Static and Dynamic Panel Data Analyses of Insurance for Escalating Climate Change

Abstract

Koohi-Kamali, Feridoon and Flaherty, Michael

This paper examines the relationship of climate shocks impact on private climate insurance by a number of econometric models. The least squares estimator produces biased estimates of such impacts because it ignores time-invariant unobservable influences; the FD estimator has a similar drawback with a more parsimonious modeling by using lagged dependent variables in first differences since the FD lagged are also endogenous. We employed a two-stage GMM dynamic Arellano and Bond (1991) model with internal lagged instruments defined in levels for endogenous FD lags. We applied these models to annual panel data sets of climate shocks of fire and water disasters and insurance prices for the US states over 2003-2013. Our best estimates of climate effects obtained from the dynamic panel model suggest the frequency of both water- and fire-based disaster events are positively and significantly related to insurance premiums over the period. In most models, the correlation with disaster incidence is strongest at the second lag, indicating that there is a reaction period after the occurrence of disaster events before they are fully absorbed in insurance prices. The dynamic panel test for the null hypothesis of second and third-order serial correlation, and that for weak instruments are both rejected. There is some suggestive evidence from descriptive data that the positive empirical relationship between environmental disasters and the price of climate insurance is likely to be driven by increasing volatility of climate shocks, lending support to the concern that the growing impact of climate change will be accompanied by increasing insurance costs and weaker private insurance coverage.

Sept 2021

Keywords: Panel Data, dynamic GMM, fire & water disaster, Insurance

JEL Classification: C25, C33, Q23, Q25

Static and Dynamic Panel Data Analyses of Insurance for Escalating Climate Change

Koohi-Kamali, Feridoon and Flaherty, Michael

Introduction

Anthropogenic climate change will significantly shift weather patterns across the globe for the centuries to come.[1] This is anticipated to increase the frequency and intensity of certain weather-related hazards and to shift the boundaries of exposed and vulnerable regions. At this point, even the most aggressive mitigation policies cannot stop climate change; decades of excessive greenhouse gas (GHG) accumulation have ensured a sort of climate change inertia, and even today's most optimistic mitigation pathways would be consistent with increased damages for decades to come.[2] Additionally, there are a number of potential tipping point approaching, which, if passed, make a "return to normalcy" impossible.[3] With significant global climate change looming on the horizon, the urgency of climate change adaptation strategies is becoming more evident.

Insurance programs may not initially spring to mind as instruments in the climate change adaptation toolkit, but these institutions can be vital for sustainable adaptation. Their support is of critical importance.[4] Insurance programs provide some amount of economic stability by spreading risk around a wide base of participants and by efficiently distributing resources to those in most dire need. In an abundant society, single events need not bankrupt individuals, households, or communities. As climate change alters the patterns of storms, droughts, floods, wildfires, and so on, the reliance on insurance systems to manage risk and loss might become ever more necessary. The climate change battle is one that demands collective action; this is very well-trodden ground when considering the necessity of global GHG reduction, but it also applies in the case of risk management and recovery. Insurance programs provides a mechanism for those latter cases; members of society pay into a pool of coverage that perhaps only a subset might ever utilize. By its very nature, an insurance program is a collective solution to relatively acute

[1] Cook et al. (2016) finds roughly 90-100 percent of publishing climate scientists agree that humans are, to a large degree, causing climate change (referred to in the paper as global warming and anthropogenic global warming (AGW)). IPCC (2014) lists the emissions of greenhouse gases (GHGs) carbon dioxide, methane, and nitrous oxide as primary causes of this climate change.

[2] The IPCC Fifth Working Group's most aggressive mitigation pathway (RCP2.6) is consistent with increased global mean surface temperatures between 0.3°C and 1.7°C and sea level rise 0.26 and 0.55 meters.

[3] Oft-discussed climate tipping points include the melting of ice sheets, the thawing of permafrost, the dying off of coral reefs, massive deforestation.

[4] Insurance is accounted for as an adaptation strategy in the IPCC Fifth Working Group, as an economic option under the "institutional" category.

problems.[5] Risk is often further diluted when insurance companies themselves insure against large loss events in the reinsurance market.

The utilization of insurance programs is likely to become more important yet more precarious as climate change shifts the risk profiles of weather events and natural hazards. Increasing occurrence of hazards, such as storms or wildfires, means a higher proportion of the population and a higher proportion of assets will be subject to increased risk of damage or loss. In other words, more people and property will be getting hurt or damaged, and insurance will face increased numbers of claims handled. Standard economic theory holds that an increase in the demand for insurance should lead to an increased price (i.e., premiums) paid to insurance companies. On the other hand, climate change damages might effectively act as a drop in income to the affected region, since, after covering damages, households and businesses have less to spend on remaining goods. Were this to be the case, it is possible to imagine a fall in demand which would counteract the upward price pressure. Covering damages while insured is preferable to covering damages uninsured, but it can be seen as a drop in income either way. At the same time, however, higher numbers of insurance payouts will likely increase costs for insurance businesses and cut into profitability, thereby putting downward pressure on the supply of such policies and programs. An increase in demand and a decrease in supply would tend to be consistent with increased prices, but it is possible that a widening pool of policyholders would limit risk and allow lower average premiums.[6]

There is an extensive literature on the drivers of insurance premiums and of insurance participation (exposure). This paper expands beyond some of the more typical socio-economic indicators to include climate-change-driven disasters as a possible explanatory variable. In Beenstock, Dickinson, and Khajuria (1988) the price of insurance (premiums) is increasing in income and the average probability of loss while ambiguously related to interest rates; at the same time, exposure is positively related to income and ambiguously related to probability of loss and interest rate. MacDonald, et al (1990) explore a hedonic model reflecting that housing price differentials between homes subject to different levels of flood vulnerability might be related to insurance premium prices; if the differential in

[5] The profit-maximizing nature of private insurance can lead to a whole host of problematic issues, which are not explored in detail in this analysis. Businesses interested in minimizing costs might reject certain insurance claims, refuse coverage for certain at-risk groups, etc. Perhaps regulation can minimize this, or government-run insurance offers a solution, see Discussion section below.

[6] It should be pointed out that the "average premium" presented in insurance literature and throughout this research is the average price paid by policyholders and does not include the $0 paid by those who have chosen to forego insurance. So, it is possible for the average premium paid by policyholders to drop while the average price of insurance across all households increases if the rate of participation increases. We will use the term *average premium* to refer to the average premium paid by a policy holder (total statewide premiums divided by total number of policies in the state), while we will use *per capita premium* to refer to premium as a proportion of state population (total statewide premiums divided by state population).

home prices between a high-risk area and a low-risk area are less than the differential in insurance rates, it might indicate that the housing market is not accurately assessing risk, and if housing price differentials are larger than insurance premium differentials this might be accounted for by considering the noninsurable household elements at risk of loss. Looking across countries, Browne, Chung, and Frees (2000) argue that barriers to trade that limit foreign insurance engagement in domestic markets, lead to less insurance consumption; This builds on Kim (1992) which relates the trade barriers rather directly to higher prices. Born and Viscusi (2006) find that while disasters and "blockbuster catastrophes" may lead firms to raise premiums in the wake of large destructive events, they can also lead insurers to exit the marketplace.

In this paper we follow a number of estimation approaches to modeling the impact of climate disasters on insurance premium; in particular we employ the dynamic panel data model, a relatively unexplored estimator in the climate disaster literature. We employ a 11-year panel data of insurance premium and climate shocks across the 50 US states that provide an improved, or at least a different, empirical strategy to elucidate the relationship between climate shocks and climate insurance prices.

Data

This study utilizes state level data for all the fifty US states, plus Washington, D.C. over the period 2003-2013; all expressed in dollar year 2017 (USD). Average state-by-state insurance premiums (i.e., the average annual price for property insurance coverage) are taken from an annual report published by the National Association of Insurance Commissioners (NAIC): "Dwelling Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report." The dataset is extensive, with information available on exposure, insurance premiums, and average premiums broken out by insurance range and by policy type. For the purposes of the current analysis, we do not distinguish between property insurance classes, choosing instead to focus on average insurance premiums for all policy classes. Data from the NAIC report have been used in analyses pertaining to the economics of disasters in recent years[7]. The NAIC data affords estimates of insurance data for all states, and its annual publication accounts for the evolution of state insurance markets annually. However, the data have some limitations. Since insurance coverage norms vary greatly by state, policies that cover earthquakes are not included in the data. Moreover, as insurance norms and regulations vary by state, there is some simplification

[7] For example, Klein (2011) considers inefficiencies in the residential property insurance markets that may lead to instances of underinsurance. Thomas, et al (2017) use the data in measuring the costs of wildfires in the United States.

taken to group coverages into consistent categories; however, this factor should not be significant in our study as we look at average premiums for *all* policies.

We employ data on types of climate disasters from the Federal Emergency Management Agency's (FEMA) "OpenFEMA Dataset: Disaster Declarations Summaries - v2". The FEMA's emergency declarations reflect cases when an emergency event overwhelms state and/or local government capabilities, and additional support is needed. The declaration follows a formazilized process whereby local leaders petition the aid of the federal government. There are three categories of FEMA declarations: emergency declarations, major disaster declarations, and fire management grants. Essentially, the major disaster declarations are used in response to events so damaging the local governments are completely overwhelmed and long-term recovery efforts might be necessary, whereas an emergency declaration is typically applied when local governments need more limited support and long-term recovery should not be necessary. The fire management (also referenced as fire suppression) grant is a new instrument aimed specifically at providing federal assistance to state and local governments handling fires; it is primarily a cost sharing program and does not have the same network of programs as major disaster or emergency declarations[8]. One of the drawbacks of FEMA declaration data usage, however, is the subjective nature through which a disaster is declared. A disaster is not declared based on some consistent, objective measure, but rather as a result of a petition and review process by local and federal leaders. In evaluating the period 2003 through 2013, we are covering six years under the George W. Bush presidency and five under the Obama administration, and it is reasonable to consider that the differences in political worldviews and the different administrators in positions of power may influence the declaration process.

Data on median household incomes is taken from the U.S. Census Bureau's Historical Income Tables (H-8: Median Income by State). In using the median income, rather than a mean, we are limiting our focus to income movements in the middle of the income distribution. We attempt to capture the relationship between income movements for the *typical* household, however, in doing so we might be missing some of the dynamics at either end of the income distribution. Often, those most vulnerable to disaster impacts are those households in more economically tenuous circumstances. Further research might consider instead per capita incomes or incomes for lower quintiles, for instance. Table 1 presents the descriptive statistics for the variables employed in this study. We note that the state

[8] In addition, there are four programs through which disasters assistance is distributed: Individuals and Households (IH) program, the Public Assistance (PA) program, and the Hazard Mitigation (HM) program. Each event is categorized as one of four disaster types: major disaster, emergency declaration, fire management, and fire suppression. For the purposes of the analysis here, we consider all disaster declaration types and all FEMA disaster programs.

components of premium and income standard deviation account mainly for the aggregate while those for water and fire shocks are somewhat lower than the time-series components.

Estimation models

The determination of the relationship between disaster occurrence and prices in the property insurance market requires estimating a number of models of the relationship between climate shocks and insurance prices. We regress insurance premiums on the lagged premium, water-related and fire-related disaster events, as well as income, for the period 2003-2013. Data is available on a state-by-state basis, and as we are dealing with 50 states (plus the District of Columbia), with 11 years of data, we consider a number of panel data and dynamic panel data models. We find the occurence of both water- and fire-based disasters are positively and significantly related with insurance premiums over the period, lending support to the concern that the growing impact of climate change will be accompanied by increasing insurance costs. In most models, the correlation with disaster incidence is strongest beginning at the second lag, indicating that there is a reaction period after the occurrence of disaster events before they are fully accounted for in insurance prices.[9]

Econometric Models

Linear Ordinary Least Squares

The ordinary least square (OLS) model assumes constant slope coefficients across all observations and *iid* error terms. The model is specified by stacking all data over *i* (cross-section) units and *t* periods into a single regression with $N*T$ observations. This pooled OLS model is given by the equation

$$y_{it} = \alpha + x_{it}\beta + u_{it} \tag{1}$$

However, with a pooled OLS model, errors are likely to be serially correlated, though can be corrected by applying robust corrected standard errors. A more significant drawback of this model is that it ignores individual-specific effects such as instances wherein coefficients are affected by certain cross-sectional characteristic differences for each observational unit. Even if the model does not exclude observable time-invariant variables, it is quite possible that there are unobservable time-invariant effects that would end up grouped into the error term. If such

[9] Increasing climate disasters may price some households out of coverage, which (a) weakens the insurance system and (b) leaves the general society more vulnerable to future climate events.

unobservable cross-sectional effects exist, the least squares coefficient estimates would be inconsistent due to omitted variable bias.

Panel Data Analysis

A more flexible model can account for change across observational units. The OLS *between estimator* is one approach to obtaining such a change over i and β_i, with time variables held constant. The between estimator, however, would not resolve the inconsistency problem that is caused by unobserved time-invariant effects, while the alternative of allowing β to change over both i and t consumes too many degrees of freedom.

Allowing each observational unit to have a different or unique intercept, while keeping the slope coefficient estimates constant across observational units is a parsimonious way to model the individual-specific effects; preferable to a model of coefficient variables that change cross-sectionally but not across time. This is the underlying concept of the fixed effect model.

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \qquad (2)$$

Here, the α_i represent the random variables that capture unobserved heterogeneity. We note that the standard panel data models are based on the assumption of strict heterogeneity, defined as:

$$E[\varepsilon_{it}|\alpha_i, x_{i1}, \ldots, x_{iT}] = 0, t = 1, \ldots, T \qquad (3)$$

This implies errors are assumed to have mean zero conditional on the past, current, and future values of explanatory variables.

There are two processes frequently used to estimate equation (2) that control for time-invariant inconsistency. The *within* estimator, also known as the fixed-effects estimator (FE), controls for time-invariant factors by an OLS estimation using *de-meaned* variables, thus removing all cross-sectional differences. However, another model with similar results that has gained greater attention is the first-difference estimator (FD). This model removes the time-invariant variables by employing variables in first-differences. The two methods produce very similar estimates, and since we later develop a dynamic panel data model employing the first-difference estimates, we confine the analysis here to the FD estimator. A general representation of the FD estimator is given as:

$$y_{it} - y_{it-1} = (x_{it} - x_{it-1})\beta + (\varepsilon_{it} - \varepsilon_{it-1}), i = 1, \ldots, N, t = 2, \ldots, T \qquad (4)$$

Since FD estimator has fewer observations on its variables than FE estimators, it is less efficient if $T > 2$; in practice, however, the outcomes tend to be similar. We note that with the panel data within-estimator, heteroskedasticity in individual errors is more important to correct at the cluster-level rather than across individuals, that is, correction requires cluster-robust rather than standard robust methods.

Dynamic Panel

Standard (or static) panel data models do not reveal important dynamic features of datasets that evolve over time. Lag values of dependent variables are determined by past covariates and typically display significant ability to predict its behavior. Excluding such lags can result in inconsistency. This is the reason for the strong assumption of strict heterogeneity in panel models. Removing the assumption of strict heterogeneity, through the inclusion of lagged dependent variables, creates a different endogeneity concern - the serial correlation of error terms will ensure correlation with lagged variables. However, consistency can be achieved in the case of weak exogeneity through the use of lagged values as instruments that are contemporaneously uncorrelated with the error terms to obtain consistent estimates. Dynamic panel data models encompass this category of models, assuming weak exogeneity which precludes only the contemporaneous correlation of instruments and regressors.

$$E[z_{is}u_{it}] = 0, s \leq t, t = 1, \dots, T \tag{5}$$

This makes that the current period error terms are uncorrelated with instruments defined in terms of the additional lagged dependent variables. One approach in the literature, outlined in Anderson and Hsiao (1981), has employed a single instrument for each endogenous variable using two-stage least squares (2SLS) estimators. This estimator is based on a model with the first differenced endogenous variable, Δy_{it-1}, instrumented by y_{it-2}, since that second lag is uncorrelated with $\Delta\varepsilon_{it-1}$, see Holtz-Eaiin 1988. A weakness of this model, however, is disregarding more lagged dependent variable instruments, which can be employed to improve efficiency. An alternate estimator is found in the over-identified case, when there are more instruments than endogenous variables, based in the general method of moments (GMM) estimation that makes the variance-covariance matrix of error terms as small as possible. Arellano and Bond (1991) offer a more efficient, GMM-based model, with endogenous lag-dependent variables formulated in first differences, instrumented by the availability of a multiplicity of lagged-dependent variables defined in levels, see also Ziliak (1997); MaCurdy (1981) for applications to labor market. This method for this dynamic first-difference estimator is known as the *two-step* GMM; it is given by:

$$\begin{aligned} (y_{i3} - y_{i2}) &= \gamma(y_{i2} - y_{i1}) + \beta'\Delta x_{i3} + \Delta v_{i3} \\ (y_{i4} - y_{i3}) &= \gamma(y_{i3} - y_{i2}) + \beta'\Delta x_{i4} + \Delta v_{i4} \\ &\dots \\ (y_{iT} - y_{iT-1}) &= \gamma(y_{iT-1} - y_{iT-2}) + \beta'\Delta x_{iT} + \Delta v_{iT} \end{aligned} \tag{6}$$

This model is dependent on the orthogonality condition (see Pesaran, 2015):

$$E(y_{it-s}, \Delta\varepsilon_{it}) = 0; t \neq s \tag{7}$$

An effective application of the Arellano-Bond estimators considers a number of other issues as well. First, the model should be tested against the static FD model using the Hausman test. The strength of the Arellano-Bond model depends on the extent of lag correlation. If the variance of the time invariant component of the error term is large relative to the time-varying component (if $\frac{\sigma_\alpha^2}{\sigma_\varepsilon^2}$ is large), or if the lag correlations are weak and close to a random walk (if $\gamma \approx 1$, then it can be shown that the levels y_{it} are only weakly related to the differences Δy_{it}), then the Arellano-Bond estimator instruments are weak, and an alternative estimator should be employed[10]. The Sargan (1985) test based on the test statistic $\chi_{SM}^2 = \frac{\hat{\beta} IV}{\hat{\sigma}^2 IV}$, discussed in Pesaran (2015), is designed to determine if Arellano-Bond instruments are weak, that is, the estimator employes more orthogonality conditions than necessary. Another important issue to consider is the dependency of the Arellano-Bond estimator on the assumption of serially uncorrelated errors. Negative first-order serial correlation is trivial and expected for the A-B model[11]. However, Arellano and Bond (1991) suggest that second order and higher serial correlation must be checked in order to assure the exclusion of higher order lag structure in the model. This model excludes any test and control of time-varying endogenous variables that are not the lagged dependent. However, the availability of internal instruments does offer a method to address any such issue, again through the use of lagged level values as instruments.

Model Specifications

OLS Model of Disaster and Insurance

We regress insurance premiums on lagged values of (*a*) the number of water or storm related disasters, (*b*) the number of fire related disasters, and (*c*) the median household income. In the pooled OLS model, all observations are stacked and regressed with one common intercept, as well as common coefficient estimates. The OLS model adopts the general format of equation (1) and is specified as:

$$prem_{it} = \alpha + \sum\beta wtr_{it-s} + \sum\lambda fire_{it-s} + \sum\gamma inc_{it-s} + \phi year_t + u_{it} \tag{8}$$

---

[10] In such cases, Arellano and Bover (1995) and Blundell and Bond (1998) suggest employing differenced instruments used for level endogenous variables as alternative estimators.

[11] Since the same instance of a (t – 1) dependent variable appears on both the left- and right hand side of the equation, in one instance subtracting the $t - 2$ dependent variable and in the other being subtracted from the time t dependent, see (6) above.

where $prem_{it}$ represents the average insurance premium in state i at time t, wtr and $fire$ quantify the number of water-related and fire-related disaster impacts respectively, inc stands for the median household income at the state level, and $t = 1,\ldots,11$. The variable $year_t$ controls for influences on premium levels derived from the passage of time. In general, we expect a positive relationship between premium price and disaster incidence. While there is some literature considering insurance coverage as an inferior good, we generally expect a positive relationship between premium and median income as well.

Panel Model of Disaster and Insurance

In order to account for any unobserved heterogeneity in the model, we run a fixed effect model with first difference estimators, based on equation (4). The entire model is expressed in first-differences (save the year variable). The best specified model which captures the significant lag structure in first differences is given by:

$$\begin{aligned} prem_{it} - prem_{it-1} &= (wtr_{it-2} - wtr_{it-3})\beta_1 + (fire_{it-2} - fire_{it-3})\beta_2 \\ &+ (fire_{it-3} - fire_{it-4})\beta_3 + (fire_{it-4} - fire_{it-5})\beta_4 \\ &+ (inc_{it-2} - inc_{it-3})\beta_5 + \phi year_t \\ &+ (\varepsilon_{it} - \varepsilon_{it-1}), i = 1,\ldots,N, t = 2,\ldots,11 \end{aligned} \tag{9}$$

Dynamic Panel Model of Disaster and Insurance

Finally, we specify the Arellano-Bond model (equation (6); we note the model and its lag structure are entirely in first differences, while instruments are in lagged levels, hence the first instrument becomes available in period $t = 3$, and each further period provides an additional instrument.

$$\begin{aligned} (prem_{i3} - prem_{i2}) &= \gamma(prem_{i2} - prem_{i1}) + \beta' \Delta x_{i3} + \phi \Delta year_3 + \Delta v_{i3} \\ (prem_{i4} - prem_{i3}) &= \gamma(prem_{i3} - prem_{i2}) + \beta' \Delta x_{i4} + \phi \Delta year_4 + \Delta v_{i4} \\ &\cdots \\ (prem_{iT} - prem_{iT-1}) &= \gamma(prem_{iT-1} - prem_{iT-2}) + \beta' \Delta x_{iT} + \phi \Delta year_T + \Delta v_{iT} \end{aligned} \tag{10}$$

where

$$\begin{aligned} \beta' \Delta x_{it} &= (wtr_{it-2} - wtr_{it-3})\beta_1 + (fire_{it-2} - fire_{it-3})\beta_2 \\ &+ (fire_{it-3} - fire_{it-4})\beta_3 + (fire_{it-4} - fire_{it-5})\beta_4 \\ &+ (inc_{it-2} - inc_{it-3})\beta_5 \end{aligned}$$

We note that we restrict the maximum number of level instruments for the autoregressive term in (10) to 5 in order to limit efficiency loss due to a large number of instruments available from this model. We estimate equation (10) three ways: first considering both water and fire disasters, then considering fire disasters alone, and finally considering water disasters alone.

An issue arises in the Arellano-Bond model when considering the potential feedback effects of climate change on the various variables in the model. Climate change might be expected to be correlated with shocks to disaster incidence, income, and any number of unobserved variables in the error term. As such, the current period error term may be correlated with future values of the explanatory variables; this represents a major endogeneity issue. The problem, however, can be corrected for through the use of further lagged levels as instrumental variables. In addition to the primary models which consider the variables exogenous, we estimated a number of specifications with disasters and/or income considered endogenous or predetermined; the exogenous model provided the best fit. Further, were there any endogeneity issues pertaining to disaster incidence and/or income, it would have theoretically persisted through the years. The fact that the model does not exhibit any higher-level serial correlation gives us confidence that the exogenous specification is appropriate. In any and all cases, applying instrumental variables, the number of instruments should be limited since, given potentially a large pool of instruments, increased utilization of the number of instruments decreases the degrees of freedom and affects efficiency.

Empirical Results

We analyze the relationship between climate-related disasters and insurance prices, first for the entire United States (fifty states, plus the District of Columbia); then confine the analysis to coastal states only: those most vulnerable to the effects of climate change (particularly water and storm related events). The panel data set is balanced, with disaster, income, and insurance premium data available for all states over the entire eleven years 2003-2013. For both data sets, a pooled OLS model was considered, followed by a first-difference model to account for cross-sectional heterogeneity, and finally three closely related Arellano-Bond dynamic panel models - one considering both fire and water related disasters, the second considering fire alone, and the third considering water alone - are run. We present the estimates for these models in Table 2 for all 50 states and Table 3 for the data set limited to coastal states; columns 6 and 7 are examined in the next section.

We find that water-related and fire-related disasters are positively and significantly correlated with increased insurance premiums. The pooled OLS (8) estimates of all 50 states and D.C. are given in the first column of Table 2; they are notably larger than the estimates derived from panel and dynamic panel models discussed below. These oversized estimates reflect the omitted variable bias inherent when cross sectional heterogeneity of states is not accounted for.

Our panel model for all states is given in the second column of Table 2. Here the first-difference estimator (9) is used to control for any time invariant factors, which would weed out the omitted variable bias that is caused by cross sectional factors. The third lag of the first-difference variable is not significant for fire-related disasters, but generally all other disaster-related explanatory variables are statistically significant.

The Arellano-Bond estimates, considering all states, are given in columns 3 through 5 of Table 2. The estimates in column 3 reflect the A-B model with all explanatory variables considered strictly exogenous, looking at both fire and water related disaster shocks, column 4 considers only fire disaster shocks, and column 5 considers only water-related disaster shocks. In all three of these specifications, all disaster-related explanatory variables are positively correlated with premiums; all are significant other than the 4th lag of the fire disaster in the jointly-specified model of column 3.

Columns 1-7 of Table 3 present the same regression estimates, however in this case the analysis is confined to states that share a border with either the Atlantic or Pacific Ocean. Of the fifty states (plus D.C.), only 23 are coastal states. There is reason to believe coastal states are particularly vulnerable to climate change impacts. Rising sea level, and the increased frequency and intensity of storms threaten increasing damage to communities in coastal areas. However, there are a number of confounding factors that make the relationship less clear. For instance, coastal areas might be high demand residential areas or they might be heavily trafficked commercial ways, and in either case it could be the case that increased infrastructure offers some amount of adaptive capacity that absorbs a part of the changing climate shocks. In any case, generally the relationships are similar in the coastal states to the full set of states, but the level of significance is much lower.

The serial correlation estimates in Table 4 report the Arellano-Bond test for serial correlation in first-differenced errors. Failure to reject the null hypothesis of no serial correlation is trivial and expected for the first order; failure to Reject the null hypothesis of no serial correlation in any higher order reflects model misspecification. In all models, we reject the null hypothesis for first order serial correlation, but we are unable to reject the null for higher orders; hence, there is no evidence of second and third order serial correlation. We also report the Sargan test of over-identifying restrictions. The null hypothesis is that the over-identifying restrictions are valid, thus a rejection of the null indicates that the model or instrumental variables used may need to be respecified[12]. The test results would indicate the model is appropriately specified and the instruments used are valid[13].

---

[12] The Sargan test is appropriate with robust standard errors are used since the test assumes a homoscedastic distribution of errors; we therefore respecify the models, minus robust errors for the explicit purpose of carrying out the Sargan test.

[13] Arellano and Bond (1991) note that the Sargan test over-rejects in the presence of heteroskedasticity. However, the conditional nature of this test should be emphasized.

Discussion

What are the public policy implications for the robust evidence of the climate shock effects on demand for private insurance presented above? Are frequency and magnitude of climate shocks driving the price of private insurance coverage out of reach for those most vulnerable to their adverse effects? There are at least three issues in the above analysis that stand out as most relevant to climate public policy insurance. First, the impact of increased climate insurance demand derives simply from wealthy consumers who chose to purchase property in desirable but prone to disaster zones. To address this question, we re-estimated the third column of Tables 2 and 3 excluding the most expensive properties from the sample. We employ two thresholds for this purpose, properties above $500,000 and those above $400,000. Since the results are very similar, we report only the former in column 6. A comparison of the estimates in both tables indicates a minor impact on climate insurance demand since the columns 3 and 6 estimates are quite similar with minor differences.

Second, the principal question is the impact the frequency and severity of climate shocks on climate insurance price. One way to quantify such an impact is counterfactually by assuming the shock effects remain unchanged over the sample period, that is we employ only the first year of the sample observations on climate shocks to estimate a restricted model. Then the predicted mean difference of the dependent variable between this regression and that obtained from the main model of column 3 is 1.12 % (the difference between the predicted values of $994 for column 3 and $983 for Column 7); this difference is too small to be credible, given the magnitude and scope of climate shocks in the United States, particularly more recently. The reason is that the expected value of the dependent variable in column 3 is taken over the sample time span that sharply changes from high to low mean values in years of severe disaster events and years of mild events, bringing the mean insurance premium price over the sample time-span close to that in the first year of the sample[14].

This points to the volatility impact of climate shocks on the pattern of change in the mean values of climate insurance demand. Table 5, top three rows, reports the increases in the variances of the frequencies of climate shocks over the earlier part of the sample time-span (2003-2007) and that of the latter part (2008-2013) are 77% for fire, 27% for water, and 88% for variance difference of fire and water. This suggest changes in insurance premium price is more severely influenced by unpredictable volatility of climate shocks; the insurance cost is driven more by change in its variance than its mean. Given data of sufficient duration the volatility of analysis

[14] For instance, the first sample descriptive mean for 2003 is 111 while those for the last three years 2011-2013 are respectively 232, 89 and 63.

would be an obvious approach to employ for the quantification of climate shocks[15]. With an 11-year sample, this option is not available to us that typically requires a sample of minimum of 250. However, the longest time-series of premium unit price available covers a 25-year period from 1993 to 2018; shown in table 5, bottom three rows. Some related descriptive data from this longer series is suggestive. To that end, we have divided the 25-year period data into five groups, each with five observations to obtain group mean and variance of climate shocks. This offers us five values of mean premium unit price and five corresponding climate shocks mean and variance values. To obtain sharper results from the states that have experienced the largest increase in the frequency of climate shocks by selected by the nine states that had experienced the highest frequency of shocks (by 103%) in both in the earliest 1993 and the latest 2018 year. Table 6 shows the correlation between logarithm of five mean premium values and logarithm of five mean and variance values for the frequencies of the shocks; since price data is not separately reported by types of climate shocks, we report the correlation for fire and water shocks combined.

First, the log mean premium price increases *monotonically* over 1993-2018; we note that prices are in log terms; hence the values indicate significant upward rise in price levels over the 1993-2018. What explains this pattern, is the pattern mainly linked to the mean frequency of the climate shocks or its variance? The correlation between log of mean unit price and log of mean shocks frequencies, though positive as expected, is at just below 50%, rather mild. By contrast, the correlation between log of price increase and variance of log of frequency shocks is positively strong, and not far below 70%. This suggests that climate insurance price increase would likely be shown to be driven mainly by the volatility of climate disaster shocks if we were able to carry out a time-series volatility analysis.

The insurers cover potential loss against unpredictable **climate shocks** risk by raising prices, and that would gradually d**rain** the effectiveness of the scheme to a point of no adaptive capacity. Therefore, a complementary public scheme is likely to play a crucial role as an adaptive climate insurance tool. This brings us to the final issue, a public insurance **for** climate disaster. **In this regard, it is helpful** look **briefly** at the assessment of a well-known public insurance scheme for climate disaster that has performed effectively as disincentive for high value property climate insurance and in supporting low income climate disaster victims. The development of such a broader insurance plan might draw inspiration from, or might even directly evolve out of, the National Flood Insurance Program (NFIP) that accounts for local hazard profiles. Browne, et al (2018) find that increased participation in

[15] Volatility analysis employs the *GARCH-in-mean* model to define the return on an asset as a function of its risk by having the variance of the series as an explanatory variable; we also note that a volatile series displays not only periods of high and changing variances but also *clustering*, that is, changing variances in those periods occur tightly close to each other, see Koohi-Kamali (2021, chapter 10).

the NFIP program across Florida is positively correlated with housing permits in non-coastal counties and negatively correlated with building permits in coastal (high-risk) counties, hence imposing additional costs on building in high-risk areas. Aerts and Botzen (2011) also confirm the NFIP flood insurance, flood zoning, and building code policies are effective tools for managing the vulnerability of land use to floods. Kousky (2018) points out that the private flood insurance market has historically been very small, the number of private policies is still dwarfed by the 5 million NFIP policies that provided climate insurance cover to **low-income** households. Most private flood insurance focuses on high-value homes; these structures have higher premiums, and are more likely to combine flood and home owner coverage in the same carrier. Hence, it is unclear how much appetite the private sector has to take on more flood risk; there is an argument for a public sector role in helping low-income families afford flood coverage. However, Kousky (2018) notes that the NFIP has amassed a huge **amount of** debt, over $20 billion as of early 2018, with no way to repay this debt in the foreseeable future. The program has always had borrowing authority from the U.S. Treasury, but no consideration was given to how this would be repaid following a catastrophic loss year. Hence, capitalization by the US Congress seems an obvious solution to its indebtedness.

Conclusion

We are beginning to experience some of the earlier impacts of global climate change, and the worse is yet to come. Mitigation effort is of the utmost importance for achieving any sort of sustainable future pathways, however adaptation effort is needed in the immediate and near term to address those impacts closest at hand.

In this paper, we examined the relationship of climate shocks impact on private climate insurance by a number of models. The least squares estimator produces biased estimates of such impacts because it ignores time-invariant unobservable influences; the FD estimator has a similar drawback with a more parsimonious modeling by using lagged dependent variables in first differences since the FD lagged are also endogenous. We also employed a two-stage GMM A-B model with internal lagged instruments defined in levels. We applied these models to panels of the US states over 2003-2013. We regard the estimates of climate effects from that the A-B model as our best and they suggest the frequency of both water- and fire-based disaster events are positively and significantly related with insurance premiums over the period. In most models, the correlation with disaster incidence is strongest beginning at the second lag, indicating that there is a reaction period after the occurrence of disaster events before they are fully accounted for in insurance prices. We tested our dynamic model estimates for second and third order serial correlation; also tested by the Sargan method for weak instruments. The tests reject the null hypothesis in both cases.

There is some suggestive evidence from descriptive data that the positive empirical relationship between environmental disasters and the price of climate insurance is likely to be driven by increasing volatility of climate shocks based on a high correlation of insurance prices with the variance of the shocks. This lends support to the concern that the growing impact of climate change will be accompanied by increasing insurance costs; public options should play a bigger role in the provision of climate disaster insurance cover to overcome the diminishing ability of private adoptive schemes.

Table 1– Descriptive Statistics 2003-2013, sample size 561 (51 states by 11 years)

Variable		**Mean**	**Std. Dev.**	**Min**	**Max**
Prem	Overall	951.83	261.13	456.53	2,246.280
	Between		244.32	557.81	1,699.404
	Within		97.79	336.82	1,498.711
Wtr	Overall	1.48	1.47	0.00	9.000
	Between		0.71	0.27	3.091
	Within		1.30	-1.61	7.390
Fire	Overall	0.99	3.72	0.00	57.000
	Between		2.25	0.00	10.818
	Within		2.97	-9.83	47.171
Income	Overall	57,352.49	8,801.25	34,026.41	82,751.100
	Between		8,425.09	41,695.57	74,318.210
	Within		2,783.44	48,458.20	68,527.300

Table 2– Estimate Results, All States (reported with t-statistic)

Variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Water (level), lag 2 yr	35.647042***						
	3.87						
Water (FD), lag 2 yr		3.0083361**	5.8879112**		5.130687**	4.15333**	
		2.43	2.62		2.13	1.91	
Fire (level), lag 2 yr	12.135086***						
	4.48						
Fire (level), lag 3 yr	3.4843661						
	0.75						
Fire (level), lag 4 yr	1.4823459						
	0.28						
Fire (FD), lag 2 yr		1.8545144***	2.4631967**	2.2641182**		2.210858**	
		3.66	2.52	2.44		2.61	
Fire (FD), lag 3 yr		1.5712862	1.5548603*	2.1753433***		1.891641**	
		1.37	1.74	3.03		2.13	
Fire (FD), lag 4 yr		2.4289783***	2.8904749*	3.4991834**		2.866617*	
		3.67	1.50	1.95		1.72	
Income (level), lag 2 yr	-.00502776***						
	-3.15						
Income (FD), lag 2 yr		0.00198562**	0.00291534**	0.00238243**	0.00083942	0.0027431**	0.00130973
		2.30	2.62	2.03	0.75	2.73	0.89
Year control	0.59417641***	8.2866489***	17.965413***	17.861091***	9.13054***	15.16855***	13.060961***
	11.65	6.24	6.71	6.03	5.08	7.11	4.97
Premium , lag 1 yr			0.60903846***	0.60724818***	0.78442451***	0.6603635***	0.69981757***
			5.30	4.76	6.32	6.32	4.35
constant		-16630.747***					
		-6.23					

Notes to Table 2:

Column 1: Pooled OLS regression (equation 8), for all 50 states plus D.C.

Column 2: Panel Data Analysis, First Difference Estimator (equation 9), for all 50 states plus D.C.

Column 3: Dynamic Panel Analysis, twostep Arellano-Bond estimator (equation 10) for both disaster categories, all 50 states plus D.C. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (45 instruments used; restricted to 5 for autoregressive premium term).

Column 4: Dynamic Panel Analysis, twostep Arellano-Bond estimator (equation 10) for fire disasters solely, all 50 states plus D.C. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (44 instruments used; restricted to 5 for autoregressive premium term).

Column 5: Dynamic Panel Analysis, twostep Arellano-Bond estimator (equation 10) for water disasters solely, all 50 states plus D.C. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (47 instruments used; restricted to 5 for autoregressive premium term).

Column 6: Dynamic Panel Analysis, same specification as Column 3. Dependent variable excludes insurance premiums for properties valued $500,000 or more.

Column 7: Dynamic Panel Analysis, same specification as Column 3. Year one disasters held constant for all time periods. This leads to a change in the mean of the dependent variable of 1.12 percent (from $994 to $983)

*Significant at 10%,** Significant at 5%,*** Significant at 1%

Table 3– Estimate Results, Coastal States (t-statistic in brackets)

Variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Water (level), lag 2	30.516617*						
	1.83						
Water (FD), lag 2		3.5507544*	8.1678572**		7.4717224*	6.1596019	
		1.42	2.28		1.87	1.48	
Fire (level), lag 2	9.0886656***						
	3.19						
Fire (level), lag 3	3.9799749						
	0.90						
Fire (level), lag 4	-1.8462142						
	-0.30						
Fire (FD), lag 2		1.6982519***	0.9278314	1.1978143*		1.0487687	
		3.05	0.71	1.66		1.01	
Fire (FD), lag 3		1.840412*	1.0059762	1.1403103		0.77185729	
		1.68	0.71	1.15		0.52	
Fire (FD), lag 4		1.9901918***	0.33235182	3.0639236		2.9939487	
		3.03	0.09	1.29		0.86	
Income (level), lag 2	-0.0118***						
	-5.78						
Income (FD), lag 2		0.00120709	0.00339325*	0.00149542	-0.00036987	0.0029388*	0.00023654
		1.03	1.64	0.92	-0.25	1.95	0.13
Year control	0.87228428***	6.2473772***	18.046618***	18.081164***	8.679616***	15.143685***	10.26949***
	11.69	3.78	7.32	6.36	3.02	5.15	3.66
Premium , lag 1 yr			0.59769622***	0.52687902***	0.71842904***	0.60018136***	0.64495393***
			8.39	8.11	6.34	5.96	4.49
Constant		-12531.844***					
		-3.78					

***Notes to Table 3*:**

Column 1: Pooled OLS regression (equation 8), for ocean-bordering states. with robust standard errors.

Column 2: Panel Data Analysis, First Difference Estimator (equation 9), for ocean-bordering states, with robust standard errors.

Column 3: Dynamic Panel Analysis, two-step Arellano-Bond estimator (equation 10) for both disaster categories, for ocean-bordering states. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (45 instruments used; restricted to 5 for autoregressive premium term).

Column 4: Dynamic Panel Analysis, two-step Arellano-Bond estimator (equation 10) for fire disasters solely, for ocean-bordering states. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (44 instruments used; restricted to 5 for autoregressive premium term).

Column 5: Dynamic Panel Analysis, two-step Arellano-Bond estimator (equation 10) for water disasters solely, for ocean-bordering states. Assumes strict exogeneity of all explanatory variables, includes one lag of the dependent variable (47 instruments used; restricted to 5 for autoregressive premium term).

Column 6: Dynamic Panel Analysis, same specification as Column 3. Dependent variable excludes insurance premiums for properties valued $500,000 or more.

Column 7: Dynamic Panel Analysis, same specification as Column 3. Year one disasters held constant for all time periods. This leads to a change in the mean of the dependent variable of 1.03 percent (from $1093 to $1081)

*Significant at 10%, ** Significant at 5%, *** Significant at 1%

Table 4– Summary of Serial Correlation & Weak Instrument Test Statistics,

	(3)	*(4)*	*(5)*
First Order Serial Correlation	-2.2547	-2.0992	-2.4188
Prob > z	*0.0242*	*0.0358*	*0.0156*
Second Order Serial Correlation	-0.14495	-0.33946	-1.4792
Prob > z	*0.8847*	*0.7343*	*0.1391*
Third Order Serial Correlation	-0.78395	-0.75311	-0.75016
Prob > z	*0.4331*	*0.4514*	*0.4532*
*Sargan**	46.34668	47.38876	49.44725
Prob > chi2	*0.016*	*0.0125*	*0.0328*
	(8)	(9)	(10)
First Order Serial Correlation	-1.9185	-1.7522	-2.0132
Prob > z	*0.055*	*0.0797*	*0.0441*
Second Order Serial Correlation	-0.38653	-0.43225	-0.87338
Prob > z	*0.6991*	*0.6656*	*0.3825*
Third Order Serial Correlation	-0.82479	-0.90939	-0.81872
Prob > z	*0.4095*	*0.3631*	*0.4129*
Sargan	20.03689	21.2079	22.6530
Prob > chi2	*0.8631*	*0.8166*	*0.9121*

Table 5– Mean and Variance of climate shock frequencies 2003-2013, and 1993-2018

Variance			
Year	Fire	Water	Combined
2003-2007	9.723	1.899	10.493
2008-2013	17.243	2.408	19.279
Pct change of 2003-13	77.30%	26.80%	83.70%
1993-2006	16.572	8.700	26.534
2007-2018	22.740	8.024	53.509
Pct change of 1993-2018	37.22%	-7.77%	101.66%

Table 6-Correlation Coefficient of log Mean Premium Unit price and log Mean and Variance Climate Disasters, 1993-2018

	ln_Premium	Variance of Premium	Disaster Number	Disaster Variance
1993-1998	6.54	5.13	1.39	0.92
1999-2003	6.65	8.09	2.84	3.31
2004-2008	6.86	4.29	3.23	3.52
2009-2013	6.98	7.83	2.87	4.30
2014-2018	7.10	5.84	2.47	3.20
correlation with ln_Premium			0.499	**0.691**

Notes: Correlation of disaster mean and variance to the log of 5-year insurance premium (unit) prices is calculated using five five-year periods. The selected subset of states comprised of Alaska, Arizona, Georgia, Kansas, Louisiana, Mississippi, Nevada, New Mexico, and South Carolina. These states have experienced the largest increase in the number of climate disaster between 1993 and 2018.

REFERENCES

Aerts, J.C.J.H. & W.J.W Botzen. “Flood-resilient waterfront development in New York City: bridging flood insurance, building codes, and flood zoning.” Annals of the New York Academy of Sciences, 1227(1), 1-82 (2011).

Anderson, T. W., and Cheng Hsiao. “Estimation of Dynamic Models with Error Components.” Journal of the American Statistical Association, vol. 76, no. 375, 1981, pp. 598–606.

Arellano, Manuel, and Stephen Bond. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” The Review of Economic Studies, vol. 58, no. 2, 1991, pp. 277–297.

Arellano, Manuel, and Olympia Bover. “Another look at the instrumental variable estimation of error-components models.” Journal of Econometrics, vol. 68, no. 1, 1995, pp. 29-51.

Beenstock, Michael, Gerry Dickinson, and Sajay Khajuria. “The Relationship between Property-Liability Insurance Premiums and Income: An International Analysis.” The Journal of Risk and Insurance, vol. 55, no. 2, 1988, pp. 259–272.

Blundell, Richard, and Stephen Bond. “Initial conditions and moment restrictions in dynamic panel data models.” Journal of Econometrics, vol. 87, 1998, pp. 115-143.

Born, Patricia, and W. Kip Viscusi. “The catastrophic effects of natural disasters on insurance markets.” Journal of Risk and Uncertainty, vol 33, 2006, pp. 55-72.

Browne, Mark J., Carolyn A. Dehring, David L. Eckles, and William D. Lastrapes. “Does National Flood Insurance Program Participation Induce Housing Development?” The Journal of Risk and Insurance, Vol. 86, No. 4, 835-859 (2019).

Browne, M., J. Chung, and E. Frees. “International Property-Liability Insurance Consumption.” Journal of Risk and Insurance, vol. 67, 2000, pp. 73-90.

Federal Emergency Management Agency (FEMA), “OpenFEMA Dataset: Disaster Declarations Summaries - v2”, data accessed: March 2019.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey Rosen. “Estimating Vector Autoregressions with Panel Data.” Econometrica, vol. 56, iss. 6, 1988, pp. 1371-1395.

Kousky, Carolyn. “Financing Flood Losses: A Discussion of the National Flood Insurance Program.” Risk Management and Insurance Review, 2018, Vol. 21, No. 1, 11-32.

Kovacevic, Raimund M. and Semmler, Willi, Poverty Traps and Disaster Insurance in a Bi-Level Decision Framework (November 4, 2019). Available at SSRN: https://ssrn.com/abstract=3515648 or http://dx.doi.org/10.2139/ssrn.3515648

Klein, Kenneth S. “When Enough Is Not Enough: Correcting Market Inefficiencies In The Purchase And Sale Of Residential Property Insurance,”18 VA. J. SOC. POL’Y & L. 345 (2011).

Koohi-Kamali, Feridoon, “Applied Advanced Econometrics”, (2021), draft text, New School University.

MacDonald, Don N., Harry L. White, Paul M. Taube, and William L. Huth. “Flood Hazard Pricing and Insurance Premium Differentials: Evidence from the Housing Market.” The Journal of Risk and Insurance 57, no. 4 (1990): 654–63. https://doi.org/10.2307/252950.

MaCurdy, T.E “An Empirical Model of Labot Supply in a a Lfe-Cycle Setting”, (1981), Journal of Political Economy, 89, 1059-1085.

Mittnik, Stefan; Semmler, Willi; Haider, Alexander. 2020. "Climate Disaster Risks—Empirics and a Multi-Phase Dynamic Model" Econometrics 8, no. 3: 33. https://doi.org/10.3390/econometrics8030033

National Association of Insurance Commissioners. 2003-2013. “Dwelling Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner’s Insurance Report.”

Pesaran, MH. Time series and panel data econometrics. Oxford University Press, Oxford, 2015.

Thomas, Douglas, David Butry, Stanley Gilbert, David Webb, and Juan Fung. “The Costs and Losses of Wildfires: A Literature Review.” Special Publication 1215. National Institute of Standards and Technology, U.S. Department of Commerce. November 2017

United States Census Bureau. “Table H-8. Median Household Income by State: 1984 to 2018.” Historical Income Tables: Households. Data Accessed August 2019.

Ziliak, J.P., “Efficient Estimation with Panel Data when Instruments are Predetermined: An Empirical Comparison of Moment-Condition Estimators”, (1997), Journal of Business and Economics Statistics, 15, 419-431.