SESM-11-1 Structural Engineering and Structural Mechanics

On the Contribution of Reconstruction Labor Wages and Material Prices to Demand Surge

by

Anna H. Olsen

Keith A. Porter

Aug. 2011

Department of Civil Environmental and Architectural Engineering
University of Colorado
Boulder, Colorado 80309-0428

Demand surge is understood to be a socio-economic phenomenon of large-scale natural disasters, most commonly explained by higher repair costs (after a large- versus small-scale disaster) resulting from higher material prices and labor wages. This study tests this explanation by developing quantitative models for the cost change of sets, or "baskets," of repairs to damage caused by Atlantic hurricanes making landfall on the mainland United States. We define six such baskets, representing the total repair cost, and material and labor components, each for a typical residential or commercial property. We collect cost data from the leading provider of these data to insurance claims adjusters in the United States, and we calculate the cost changes from July to January for nine Atlantic hurricane seasons at fifty-two cities on the Atlantic and Gulf Coasts. The data show that: changes in labor costs drive the changes in total repair costs; cost changes can vary significantly by geographic region and year; and cost changes for the residential basket of repairs are more volatile than the cost changes for the commercial basket. We then propose a series of multilevel regression models to predict the cost changes by considering several combinations of the following explanatory variables: the largest gradient wind speed at a city in a hurricane season; the number of tropical storms in a hurricane season whose center passes within 200 km of a city; and cost changes in the first two quarters of the year. We also allow the coefficients of the regression model to be stochastic, varying across groups defined by region of the Southeastern United States and year. Our best models predict that, for any city on the Gulf or Atlantic Coasts in any hurricane season, the residential total repair cost changes vary from 0.01 to 0.25, depending on the wind speed and number of storms, with an uncertainty of 0.1 (two standard errors of prediction) given the wind speed and number of storms. The commercial total repair cost changes vary from 0.005 to 0.15 with an uncertainty of 0.08. Our models including wind speed, the number of storms affecting a city, and cost changes in the first half of the year explain roughly half of the observed variability in cost changes. Additional explanatory variables that we have not considered may account for the remaining variability. Given these models, however, there is still considerable uncertainty in their predictions. This uncertainty arises from variations between groups defined by region and year, not from variations within a given region and year.

Contents

List of Figures

List of Tables

1 Introduction

Demand surge is understood to be a socio-economic phenomenon of large-scale natural disasters in which repair costs rise, locally and temporarily, through any of several possible demand-related mechanisms. As a result, the repair cost for a property damaged in a large event exceeds the repair cost for a similar property damaged in a small event, given the same damage at the two properties. Increased repair costs after past large-scale natural disasters have been reported in the range of 20 to 50 percent. For example, after the 1994 Northridge Earthquake, insurers observed a 20 percent increase in the costs to settle claims (Kuzak and Larsen, 2005, p. 113). Commercial catastrophe modelers estimated demand surge of 10 to 40 percent after Hurricane Katrina (Guy Carpenter, 2005). The Australian Securities & Investments Commission (2007) reported that insurers saw reconstruction costs increase 50 percent after Cyclone Larry in 2006.

Institutions that indemnify properties exposed to natural disasters, such as insurers, reinsurers, and governments, pay billions of U.S. dollars in claims after large-scale natural disasters; these payments can be even larger as a result of demand surge. Anticipating whether demand surge is 20 percent versus 30 percent versus 50 percent can affect how insurers, reinsurers, and governments plan for and respond to large-scale natural disasters. Understanding the socio-economic mechanisms of demand surge should result in better predictive models and is presumably a prerequisite to controlling, and possibly reducing, its magnitude after future disasters. This work studies two of several possible demand-surge mechanisms.

In previous work, we found that there is no standard, precise definition of demand surge (Olsen and Porter, 2010). We surveyed the literature on demand surge and identified four types of definitions.

1. Demand surge is the temporary increase in local construction labor wages, material prices, and other specific costs after a large-scale natural disaster. Different references identified different specific costs. These costs included: increased construction contractor overhead and profit; possible overpayment by insurers because they could not verify all claims or they faced political pressure to pay claims quickly; payments for damage not explicitly covered in an insurance policy; and special repairs required to maintain compliance with current building codes.

2. Demand surge is the temporary, local increase in construction labor wages and material prices after a large-scale natural disaster. Here, demand surge is limited to material prices and labor wages, excluding the additional costs included in the first definition.

3. There is a general increase of reconstruction costs after a large-scale natural disaster, but no explanations are provided for the increase. In other words, increased costs result from an unidentified and unexplained, underlying phenomenon.

4. Demand surge is the discrepancy between actual monetary losses and expected (or modeled) losses.

We do not favor any of these definitions, and we do not offer a specific definition of demand surge. The existing evidence for demand surge does not establish it as a clearly delineated phenomenon. Rather, we understand demand surge to be a collection of specific, but as yet vaguely defined, socio-economic phenomena of large-scale natural disasters that result in increased costs to repair an insured property.

We describe the socio-economic phenomena contributing to demand surge as "vaguely defined" because each definition listed above and found in the demand surge literature is unclear. For example, if a temporary increase in local labor wages is a real mechanism of demand surge, how far does "local" extend, how long does "temporary" last, which construction trades must be considered, and what governs the rise of wages? A clear definition for wage-driven demand surge must at minimum address these questions.

In addition to identifying several definitions of demand surge, we proposed seven general descriptions of socio-economic mechanisms that could result in increased repair costs after large-scale versus small-scale natural disasters (Olsen and Porter, 2010). These reasons for demand surge followed from studies of the reconstruction periods after historical natural disasters, including earthquakes, hailstorms, cyclones, flooding, and wildfires, from the fourteenth, nineteenth, and twentieth centuries through the present day, in Australia, the United States, the United Kingdom, and continental Europe. Although the circumstances contributing to increased construction costs were unique to each disaster, there were common reasons for demand surge when we considered the disasters together. Our seven possible explanations for demand surge are, briefly: (1) the total amount of repair work; (2) the costs of reconstruction materials, labor, and equipment; (3) reconstruction timing; (4) construction contractor fees; (5) general economic conditions; (6) insurance claims handling; and (7) decisions of an insurance company.

Various groups have developed quantitative models for demand surge. We surveyed these models in Olsen and Porter (2010) and updated our summary in Olsen and Porter (2011). The primary developers of demand surge models are commercial catastrophe modelers, such as AIR Worldwide, RMS, and EQECAT. These companies provide quantitative estimates of risk usually resulting from natural hazards for a single property or set of properties. Their clients are mostly insurance and reinsurance companies, which are interested in anticipating their monetary losses in the event of a natural disaster. Although these models are proprietary, the standard model of demand surge multiplies the ground-up loss by a factor, typically between 1.0 and 1.6. (The ground-up loss is the calculated loss at a property before applying insurance deductibles, co-payments, or limits.) This multiplicative factor can be based on the expected loss to the insurance industry as a whole, the affected region, the type of peril, the type of property, or some combinations of these.

Researchers have repurposed existing, regional economic models or developed index-based models to estimate demand surge. Regional economic models identify and model distinct sectors of the economy, interactions between the sectors, and the interactions between economic sectors in different geographic regions. The two most common types of economic models applied to natural disaster research are input-output and computable general equilibrium models. (See, for example, Rose (2004) and Okuyama (2007) for descriptions of these models in the context of natural disasters, and see Hallegatte (2008) for an estimate of demand surge after Hurricane Katrina derived from an input-output model.)

Researchers at Florida International University developed a demand surge module for the Florida Public Hurricane Loss Model. According to their 2009 submission to the Florida Commission on Hurricane Loss Projection Methodology, the module applies “weighted average demand surge factors” to the loss from each event in the stochastic event set (International Hurricane Research Center, 2009). The model assumed demand surge is affected by the type of insurance coverage, the location of the property within Florida, and the modeled statewide loss without demand surge. In developing the model, the researchers compared the values of a construction cost index after several hurricanes to inferred values of the index had no hurricanes occurred. They assumed the discrepancies were entirely attributable to demand surge. All existing demand surge models are difficult to independently verify and validate because the full models are not published or because the underlying data are not publicly available or shown with the model.

The work described in this report addresses part of our second proposed explanation for demand surge; we study the costs of reconstruction material prices and labor wages before and after Atlantic hurricane seasons. Increased costs of reconstruction materials and labor is the most common explanation for demand surge, and therefore, among our proposed explanations, it is a good first choice for an in-depth study. Also, data are available on material prices and labor wages, whereas data pertaining to the other demand surge explanations are not as readily available.

Section 2 describes the hazard, material price, and labor wage data we use. Section 3 develops quantitative models for the cost change of baskets of repairs along the Atlantic and Gulf Coasts during nine Atlantic hurricane seasons. This section plots the data, proposes several quantitative models for the cost changes, and evaluates these models. We end this study with a discussion of our results (Section 4) and a summary of our findings (Section 5).

2 Cost and Hazard Data

To establish how reconstruction material prices and labor wages contribute to demand surge, ideally a researcher would examine economic data on daily, weekly, or monthly time scales and at a geographic resolution fine enough to distinguish the region with disaster-induced damage from nearby regions without damage. With this information one could show how costs change from before an event to immediately after and through the reconstruction phase, while learning over what geographic area these costs change. By coupling these data with characterizations of the event (for example, wind speeds, flood depths, earthquake intensities), one might find an appropriate way to distinguish natural disasters that cause demand surge versus those that do not, or in other words, the intensity and extent of excitation large enough to induce demand surge. The data just described, however, seem to be unavailable.

The data we use are from two sources: Xactimate (labor and material costs) and the National Hurricane Center (hurricane tracks and intensities). The data on materials and labor costs from Xactimate are widely used by claims adjusters in the United States. Although we are limited by the frequency, locations, and choice of materials and labor collected by others, the use of existing data allows us to compare prices and wages across many historical disasters. In Section 3 we use observations of tropical storm tracks and intensities from the National Hurricane Center in conjunction with the Xactimate data.

The economic data used in our analyses were manually entered into databases. They are not electronic copies of the original source. We checked the data for unusual entries and made corrections as necessary.

2.1 Xactimate

Xactimate is computer software developed and distributed by Xactware Solutions, a company based in Orem, Utah. Xactware collects pricing data on building repairs, contents replacement, and cleaning costs for more than 470 cities in the United States, Canada, United Kingdom, and Ireland (Xactware, 2011e,d). (At the time of this study, Xactware did not publish price lists for the United Kingdom and Ireland.) Xactware packages the pricing data into software appropriate for different applications; Xactimate is the software for estimating claims on property insurance policies(Xactware, 2011c). According to Xactware: “80 percent of insurance-repair contractors and 16 of the top

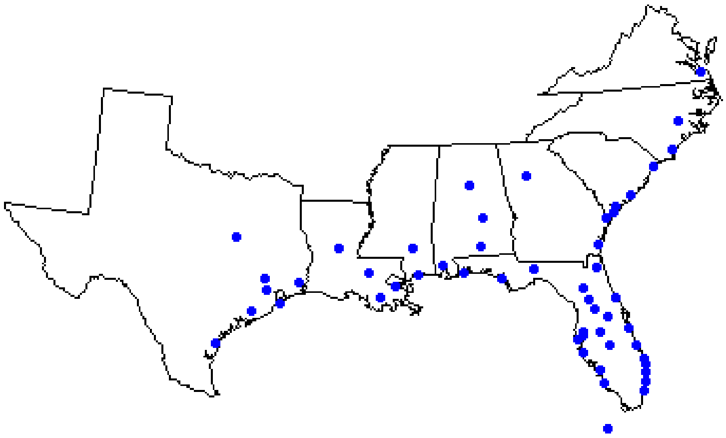

Figure 1: Cities on the Atlantic and Gulf Coasts with data from Xactimate

20—including 9 of the top 10—property insurers use Xactimate to calculate the cost of repairs" (Xactware, 2011a). Although Xactware was founded in 1986, only data since 2002 are readily available in current versions of Xactimate. Data in the years 2002 through 2008 are published at the beginning of each quarter, and data since 2009 are published at the beginning of each month.

Xactware uses the following resources to populate its price lists: "thousands of in-field estimates that are submitted to Xactware every day (i.e., estimates actually used to settle claims); market surveys of industry professionals; retail pricing research; unit-price research based on surveys with over 100,000 contractors, insurance carriers, and independent adjusters; pricing feedback from in-field users; independent pricing verification requests; customer-specific cost data; catastrophe-specific pricing research; additional research surveys; multiple third-party sources for data such as workers comp[ensation], federal taxes, state taxes, local taxes, and so on" (Xactware, 2011f). After reviewing the collected data for quality, Xactware "perform[s] a proprietary cluster analysis on various subsets of that research to identify the mid-range market price points"[1] (Xactware, 2011f). In other words, the price Xactware publishes is intended to be at the center of its observed prices for a particular item in a given city at the time of publication.

Xactware also cautions its customers to use their knowledge of "local market conditions" when applying the pricing data: "Xactware makes every effort to ensure pricing information ... represents market costs at the time of publication. Since actual market prices can vary and may change rapidly, and since many factors can affect the cost of a project (including—but not limited to—labor, equipment, and material costs as well as the rates and application of sales tax), we strongly recommend customers monitor their local markets for any such changes and adjust their estimate pricing as deemed appropriate" (Xactware, 2011b).

We collect a subset of the data Xactimate provides. We select 53 United States cities either on the Atlantic or Gulf of Mexico Coast or in Iowa. (Figure 1 maps the cities on the Atlantic and Gulf Coasts.) Data from cities near the Coasts are used to track costs during the 2002 through 2010 Atlantic hurricane seasons. For the purposes of this study, the city in Iowa represents a location outside the area of study but still in the United States. We collect data from the first quarter of 2002 through January 2011.

One of us (Porter) chose sets of repair line-items typically performed in reconstruction after natural disasters, based on his experience and judgment. The sum of the costs for the items in each set forms our basket of repair costs. For residential properties this set is: 300 square feet (sf) of composition shingle roofing; 100 sf clay tile roofing; 40 sf reglazing; 100 sf installed drywall; 100 sf carpet; and 200 sf sealed and painted wood siding. For commercial properties this set is: 100 sf 3-ply roofing; 100 sf gravel ballast roofing; 100 sf PVC membrane roofing; 100 sf metal roofing; 100 sf glass curtain wall; 100 sf installed drywall; 100 sf carpet; and 100 sf sealed and painted wood siding. Appendix A reproduces Xactimate's descriptions of these items as of July 2010. We have not compared these baskets of repair costs to actual insurance claims. We assume our baskets are consistent with the types of repairs needed after natural disasters, and thus we assume our baskets of repair costs are highly correlated with insurance claims following natural disasters.

For each repair line item, Xactimate provides several types of costs. The total cost to make a repair is the "remove & replace" cost, which is the sum of the removal and replacement costs. For each of these two tasks, the cost is

[1] Note that the "cluster analysis" Xactware performs is not the same as the statistical data analysis known as "cluster analysis." In the field of statistics, a cluster analysis identifies subsets of observations that are more alike than different, clustering them into previously-unidentified groups. (See, for example, Johnson and Wichern (2002, Chapter 12).)

the sum of: labor, material, and equipment costs; market conditions; labor burden; and taxes.[2] We employ three of these costs: remove & replace, material, and the sum of regular labor and labor burden. Since we have distinct sets of repairs for residential and commercial properties, we track a total of six baskets of repair costs.

Xactimate reports "hard costs" and "market costs." The market costs represent costs reported to Xactware in estimates of repair costs used to settle insurance claims (Xactware, 2009). Xactware also collects data on the same costs directly from construction materials and equipment suppliers as well as construction contractors. These are the "hard costs." In its documentation, Xactware explains that market costs can exceed hard costs for several reasons. For example, demand for materials or labor exceeds the supply in a city, or there are additional costs for repairs reported in estimates of repair costs but not reported by materials or equipment suppliers or people familiar with the construction labor market. In the work described in this report, we consider both hard and market costs.

The Xactimate databases are not complete for all cities at all times. This is a particular problem for many Florida cities in October 2004. We linearly interpolate missing data between the preceding and succeeding quarter or month, and we indicate when the data have been interpolated. We expect that a linear interpolation under-estimates the actual price or wage, since there must have been highly unusual conditions if Xactware was not able to use its routine procedures to report price and wage data.

2.2 Hurricane Data

For each city where we have data from Xactimate, we find the latitude and longitude from the National Atlas of the United States (2003). We presume that each city's coordinates are near its city center, and we do not use any additional information about the city, such as population, incorporated area, or metropolitan area.

The National Hurricane Center provides observations of tropical storms, including latitude, longitude, maximum wind speed, minimum central pressure, etc. (Specifically, we use the Extended Best Track File (Demuth *et al.*, 2006).) For each city where we have price and wage data, we first identify the storms that come within 200 km of the city in a given hurricane season. We use 200 km because it seems small enough to exclude hurricanes not affecting a city but large enough to allow for multiple hurricanes affecting a city in a hurricane season. Then we apply the Holland Wind Profile (Holland, 1980) at each observed point on the tropical storm's path, calculating the gradient wind speed at a city resulting from each observation at a known distance from the city center.[3]

The Holland Wind Profile finds the gradient wind speed, V_g, at a distance, r, from the center of a hurricane. The profile is:

$$V_g = \left[\frac{AB\left(p_n - p_c\right)\exp\left(-A/r^B\right)}{\rho r^B} + \frac{r^2 f^2}{4}\right]^{1/2} - \frac{rf}{2},$$

where $A = \left(R_w\right)^B$, $B = \rho e \left(V_m\right)^2 / \left(p_n - p_c\right)$, e is the base of natural logarithms, p_n is the ambient pressure (pressure of the outer closed isobar in the Extended Best Track File), p_c is the central pressure, $\rho = 1.15$ kg/m^3, $f = 2\Omega \sin(\text{latitude})$ is the Coriolis frequency, $\Omega = 7.2921 \cdot 10^{-5}$ rad/s, R_w is the radius of maximum wind speed, and V_m is the maximum wind speed.

We report the maximum gradient wind speed at a city among those calculated for each observation on a tropical storm's path, and we define the surface wind speed as 80 percent of the maximum gradient wind speed. If there is more than one tropical storm within 200 km of a city in a hurricane season, we report the maximum gradient wind speed among the maximum wind speeds of the individual storms. We use wind speed as a proxy for damage at a city: data on damage at so many cities after so many tropical storms are not available, and because our analyses are exploratory, we cannot justify using such detailed information.

We also count the number of tropical storms affecting a city in a particular time period. We define a "storm affecting a city," or a "proximate storm," to be a tropical storm that comes within 200 km of the city irrespective of any wind speed.

3 Changes in Xactimate Cost Data

In this section we examine the relationship between cost changes in the Xactimate data and observed tropical storms during the 2002 through 2010 Atlantic hurricane seasons. We first plot the data to develop a qualitative understanding of possible explanatory variables for repair cost changes. Then we propose a series of quantitative models based on our available data and our initial understanding. Finally we identify and discuss the best predictive models among those we propose.

[2] Xactimate price lists use labor costs for about two dozen types of labor, and Xactimate selects the appropriate type for a particular line item (Xactware, 2008, p. 2). The labor burden category is a markup additional to the regular labor cost resulting from any labor costs unique to that repair item or to account for the overhead to employ laborers.

[3] The gradient wind speed is the wind speed above the hurricane boundary layer, or the wind speed in the absence of friction caused by interaction with the ground or water surface. It is the wind speed 1–2 km above the surface (J. Done, personal communication, April 2011).

The questions one can address with these data should help build a foundational knowledge for demand surge. We first examine whether cost changes after a large tropical storm exceed the cost changes one would expect in the absence of such an event. In other words, does there appear to be a cause-and-effect relationship between hurricanes and unusually large cost changes? If so, we might find a relationship between the intensity of the storm and the resulting cost changes. Do more intense hurricanes cause larger changes in repair costs? From our previous work and the work of others, we expect that the intensity of a tropical storm and the number affecting a specific city would help explain large cost changes. Additional factors on the regional level or at the time of a hurricane season—such as municipal or state restrictions on labor movement or a pre-existing demand for new construction—might explain why one city damaged in a large tropical storm has larger cost changes than a similarly damaged city at a different time or place. We return to these issues in Section 4 after considering the data and developing quantitative models.

As described in Sections 2.1 and 2.2, we process the data from Xactimate and the National Hurricane Center. We develop several baskets of repair costs, specifically, baskets for the (a) total remove & replace costs, R&R, (b) labor component of repair costs, LAB, and (c) material component of repair costs, MAT, each for residential (RES) and commercial (COM) properties. We define a basket cost change, D, as the difference between the final and initial costs normalized by the initial cost. As an equation,

$$\text{cost change} = D \stackrel{\text{def}}{=} \frac{(\text{final cost}) - (\text{initial cost})}{\text{initial cost}}.$$

We know the change in the cost of these six baskets during the following time periods: July to the following January, Δ, (roughly from before to after the Atlantic hurricane season) and incrementally by quarter: from January to April, Δ_1; April to July, Δ_2; July to October; and October to January. Xactimate provides quarterly data in the years 2002 through 2008 and monthly data afterwards. In order to use data from as many Atlantic hurricane seasons as possible, we do not consider the monthly cost changes, only the cost changes reported at the beginning of one quarter to the beginning of another quarter.

For simplicity, we consider only two characteristics of the tropical storms in a given Atlantic hurricane season: the intensity of each storm and the number affecting a city. A gradient wind speed at the city center is calculated for each storm that passes within 200 km of a city, and if there is more than one such storm, the reported gradient wind speed is the largest among the storms.[4] A surface wind speed is approximated as 80 percent of the gradient wind speed. We use surface wind speeds in the data exploration (Section 3.1) and gradient wind speeds in the quantitative models (Section 3.3). We also count the number of storms passing within 200 km of the city, which we call "proximate storms."

Our choice for calculating wind speeds at the cities results in relatively small values. We do not use observations from instruments at the cities or in their metropolitan areas. Rather, the wind speed at a city is inferred from observations of the storm system. We expect, however, that the wind speeds we use are highly correlated with observations at the cities.

3.1 Data Exploration

Figures 2 and 3 show the hard and market cost changes, respectively, from July to the following January of the baskets of repair costs as functions of the surface wind speed of a proximate storm. Data plotted where the wind speed is "None" are for cities with no proximate storm. The data points are colored by year, and the shapes of the symbols correspond to the state in which the city is located. Wind speeds corresponding to Category 1 on the Saffir-Simpson Hurricane Wind Scale are shaded. To be clear, the tropical storm generating a point in these figures could have any designation on the Saffir-Simpson Scale at the time of its closest approach to the city. The Category 1 shading is intended as a reference or comparison point for the calculated surface wind speeds, not a categorization of the storm systems producing the data.

Several observations can be made by comparing the subplots within Figure 2.

- For the remove & replace and labor component baskets, the maximum cost change when there is no proximate storm is similar to the maximum cost change when there is at least one storm with surface wind speed less than 50 km/hr. In other words, remove & replace and labor component costs can increase by 10–20 percent in the absence of a tropical storm during a hurricane season. These cost changes exceed 20 percent only if there is at least one sufficiently intense tropical storm. The range of material component cost changes when there is no proximate storm is similar to the range when there is at least one storm, independent of wind speed.

- The residential baskets have more extreme cost changes (that is, a larger range of values) than the corresponding commercial baskets. The costs of residential reconstruction appear more volatile during Atlantic hurricane seasons than the costs of commercial reconstruction.

[4] Recall that we use wind speed as a proxy for damage. We assume that the storm with the largest wind speed causes all damage in a given hurricane season. Since we do not have cost data after each storm, we cannot attribute incremental cost increases within a season to particular storms.

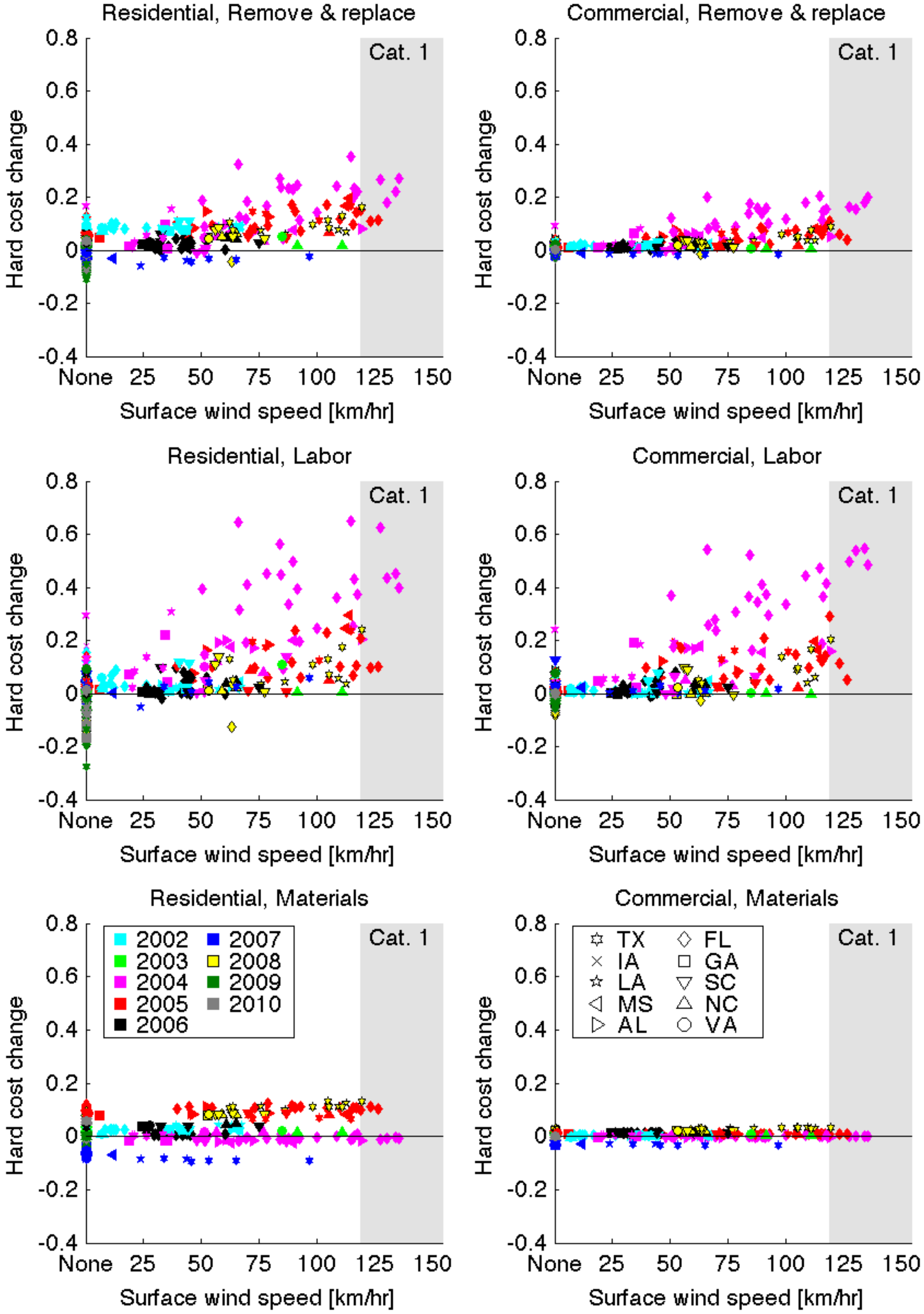

Figure 2: Observed hard cost change from July to January versus wind speed. Symbol colors represent different Atlantic hurricane seasons, and symbol shapes represent different U.S. states.

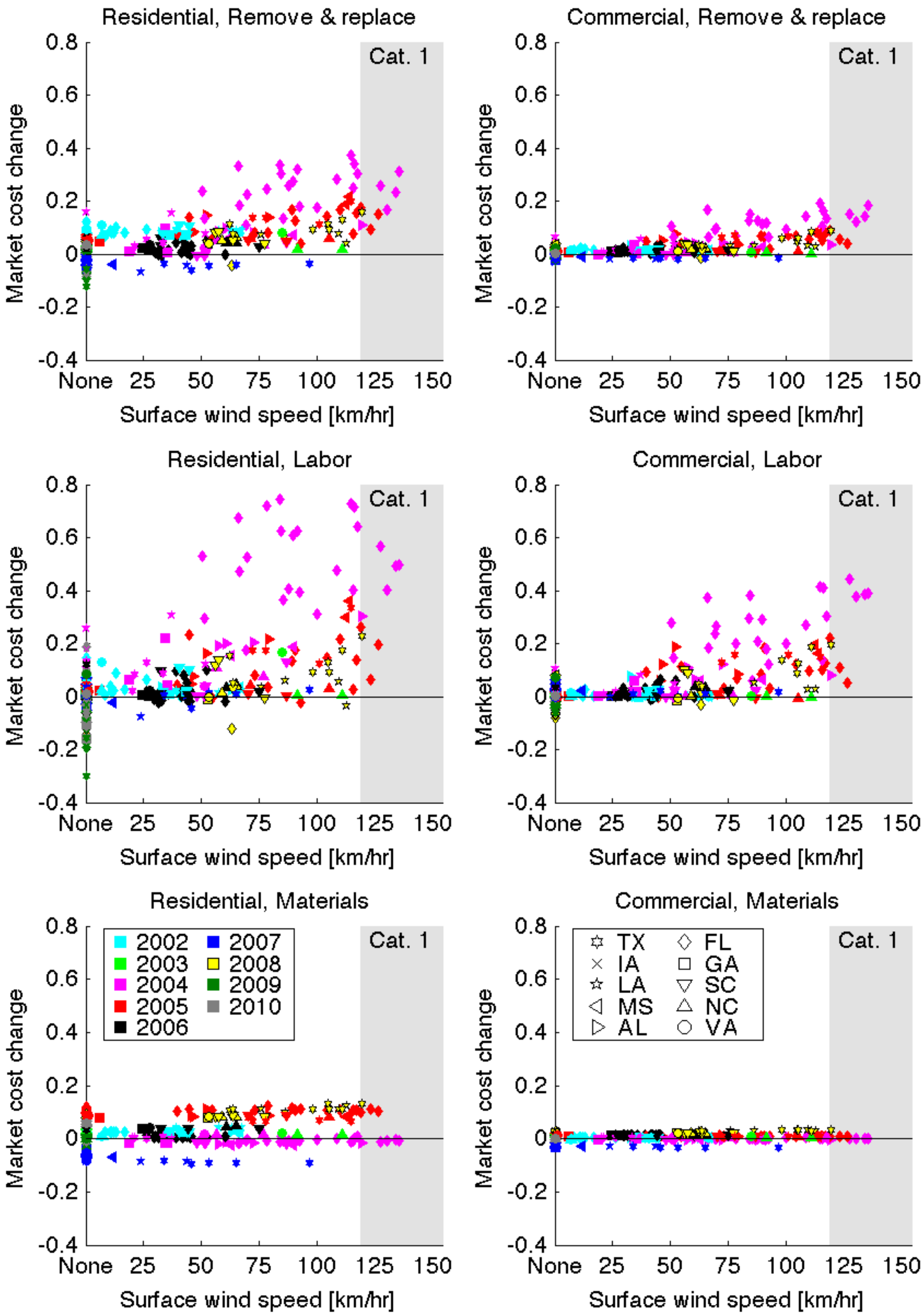

Figure 3: Observed market cost change from July to January versus wind speed. Symbol colors represent different Atlantic hurricane seasons, and symbol shapes represent different U.S. states.

- The remove & replace cost changes are driven by the labor, not the material, component. The material cost changes are well within ± 20 percent for residential properties and well within ± 10 percent for commercial properties irrespective of a proximate storm. There seems to be little correlation between wind speed and the change in cost of a material component basket. The pattern of labor cost changes is similar to the remove & replace cost changes: there is a positive correlation between cost change and wind speed for both basket types. However, for both residential and commercial, the labor component baskets show more extreme, both positive and negative, cost changes than the remove & replace baskets.

- The most extreme positive cost changes of the remove & replace and the labor component baskets are exclusively from Florida in 2004. Observed cost changes in other regions and years are never as large, implying that the circumstances causing cost changes in Florida in 2004 were unique within the geographic regions and years we studied.

Finding 1 *There seems to be a threshold wind speed above which remove & replace and labor component cost changes can exceed 20 percent. Below this threshold, there are no observations of cost changes greater than 20 percent. The threshold value surely depends on the specific metric for wind speed. Using our definition, this threshold is a surface wind speed of 50 km/hr.*

Finding 2 *The cost change of the material component baskets appears to be much less sensitive than the other baskets to wind speed. A prediction of the material component cost change would not be much different given there is at least one proximate storm, with any wind speed, than a prediction given there is no proximate storm.*

Finding 3 *The costs of our residential baskets are more volatile than the costs of our commercial baskets during Atlantic hurricane seasons. Based on these observations, we would expect more extreme, both positive and negative, cost changes of the residential, compared to the commercial, baskets.*

Finding 4 *The remove & replace cost changes are driven by the labor component, not the material component.*

Finding 5 *The most extreme positive cost changes of the remove & replace and the labor component baskets are exclusively from Florida in 2004.*

The only obvious differences between hard cost changes (Figure 2) and the corresponding market cost changes (Figure 3) are between the labor component baskets. The residential labor component basket has more extremely positive market cost changes than the corresponding hard cost changes. For example, note the greater number of market cost changes exceeding 40 percent compared to hard cost changes exceeding 40 percent. Curiously, however, the commercial labor component basket has more extremely positive hard cost changes than the corresponding market cost changes. For example, note the greater number of hard cost changes exceeding 40 percent compared to market cost changes exceeding 40 percent. The pattern of market cost changes is similar to that of the hard cost changes, although the specific values differ.

Finding 6 *(Recall that Xactware collects labor and material costs, "hard costs," directly from construction contractors and material suppliers. "Market costs" are labor and material costs derived from as-written estimates of repair costs for damaged properties reported to Xactware.) Only the labor component cost changes show a clear difference between hard and market cost changes. The market cost changes of the residential labor component basket tend to be larger than the hard cost changes. This observation is reversed for the commercial labor component basket: the hard cost changes tend to be larger than the market cost changes.*

Throughout the remainder of this section, we consider only the hard cost changes and not explicitly the market cost changes. The following statements we make about hard costs can apply to market costs as well, with adjustments of particular values as appropriate. Corresponding plots for the market cost changes are in Appendix B.

Figures 4 and B.1 show the data from Figures 2 and 3 as box-and-whisker plots instead of scatter plots. We represent the data for no proximate storm with a box-and-whisker at "None" and data for at least one proximate storm with three box-and-whiskers according to ranges of wind speed. In these plots—as in all box-and-whiskers in this report—we define an outlier to be a value greater than $q_3 + 1.5\,(q_3 - q_1)$ or less than $q_1 - 1.5\,(q_3 - q_1)$, where q_1 and q_3 are the twenty-fifth and seventy-fifth percentiles, respectively (MathWorks, 2010a).

Some of the observations associated with Figures 2 and 3 are more obvious in Figures 4 and B.1, but we make further observations about the location and spread of the data by these categories. For five of the baskets—all but the commercial material component—the median cost change increases with increasing wind speed. On the vertical scale shown, the trend in median cost change of the commercial material component cannot be discerned. For four baskets—residential and commercial, remove & replace and labor component—the inter-quartile ranges also increase with increasing wind speed. Any difference in the range of data by category is not obvious. Generally speaking, there are fewer outliers at wind speeds above 50 km/hr than at lower wind speeds. The outliers when there are no

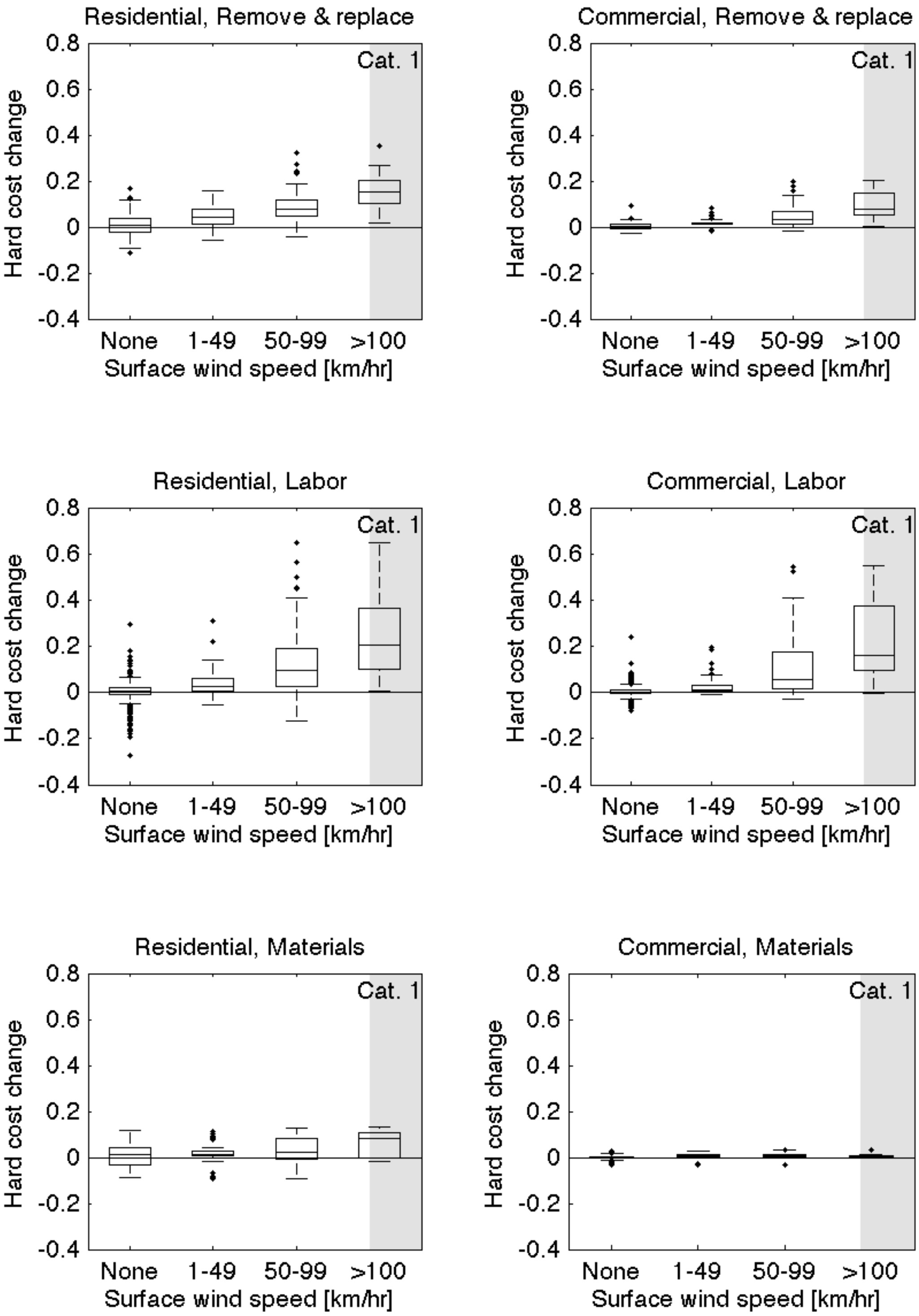

Figure 4: Observed hard cost change from July to January categorized by wind speed. The shaded wind speeds labeled "Cat. 1" correspond to a Category 1 hurricane on the Saffir-Simpson Scale.

	p-value						
Basket	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*
RES R&R	0.3232	0.3842	0.0020	0.0000	0.4497	0.0958	0.0000
COM R&R	0.2742	0.9290	0.0001	0.0000	0.4794	0.0438	0.0001
RES LAB	0.2150	0.6694	0.0020	0.0000	0.3801	0.0655	0.0001
COM LAB	0.4254	0.1536	0.0001	0.0000	0.1274	0.7566	0.0001
RES MAT	0.1120	0.0800	0.0038	0.5170	0.3417	0.0025	0.0006
COM MAT	0.0236	0.0020	0.0173	0.9903	0.2947	0.0025	0.0006

Table 1: P-values from hypothesis tests that observed hard cost changes in a given year do not differ between cities affected and unaffected by large proximate storms

proximate storms are positive and negative; outliers when there is a proximate storm are almost always positive. In summary, for these four baskets of repair costs, the distributions of the data change with wind speed. For the material component baskets, there is no clear pattern to the changes in the distribution of data with wind speed.

Finding 7 *For the remove & replace and labor component baskets, both the center and the spread of the distributions of cost change increase with increasing wind speed. For the material component baskets, there is no clear pattern to the distributions of cost change with wind speed.*

Figures 5 and B.2 categorize the data from Figures 2 and 3 by the number of proximate storms. For four baskets—residential and commercial, remove & replace and labor component—the median cost change increases with an increasing number of proximate storms. The distribution of data changes with the number of proximate storms, but there is no clear pattern to how the distributions change. For the material component baskets, there are some differences in the distributions of cost change by number of proximate storms, but there is no apparent pattern to these differences.

Finding 8 *For the remove & replace and labor component baskets, the center of the cost change distributions increase with an increasing number of proximate storms, but there is no discernible pattern to the change in spread of the distributions with the number of storms. For the material component baskets, there is no pattern in either the center or spread as a function of the number of proximate storms.*

Figures 6 and B.3 re-categorize the data by year and whether there was one or more proximate storm with wind speed greater than 50 km/hr in that year. The colors in Figures 6 and B.3 are consistent with those in Figures 2 and 3. We plot the data in this way to learn: for a given year, do the cost changes in cities with a large proximate storm significantly differ from the cost changes in cities unaffected by a large proximate storm? We use a two-sample Kolmogorov-Smirnov test of the hypothesis that data from unaffected cities in a given year are drawn from the same underlying population as data from affected cities in the same year (MathWorks, 2010b). If we cannot reject this hypothesis at a 5 percent level, we place a rectangle around the box-and-whiskers representing these data, indicating that the cost changes in unaffected and affected cities in that year may be drawn from the same underlying population of cost changes. Tables 1 and B.1 show the p-values for all tests on the hard and market costs, respectively.

In some years, there is no significant difference between cost changes in unaffected and affected cities, but in other years, there is a significant difference. For the remove & replace baskets, there is no significant difference in cost changes between unaffected and affected cities in 2002, 2003, and 2006 (and, for the residential basket, 2007). Similarly, for the labor component baskets, there is no significant difference in cost changes between unaffected and affected cities in 2002, 2003, 2006, and 2007. For the material component baskets, there is no significant difference in cost changes between unaffected and affected cities in 2005 and 2006 (and, for the residential basket, 2002 and 2003). This suggests an underlying explanatory variable—such as economic conditions immediately before the hurricane season—for which year is a proxy; in some years, the presence of a large proximate storm significantly affects cost changes, but in other years, a large proximate storm may not affect cost changes.

Finding 9 *In some years between 2002 and 2010, there are significant differences between cost changes in cities affected by at least one large proximate storm and cities unaffected by such a storm. These years are generally consistent across the six baskets of cost changes. Presumably, the year is a proxy for some underlying economic phenomenon that determines whether large proximate storms will significantly change costs in an affected city that year.*

Figures 7 and B.4 re-categorize the data according to the state where each city is located and whether there is at least one proximate storm with wind speed greater than 50 km/hr. The symbol shapes representing outliers are consistent with the shapes in Figures 2 and 3. We perform another round of two-sample Kolmogorov-Smirnov tests,

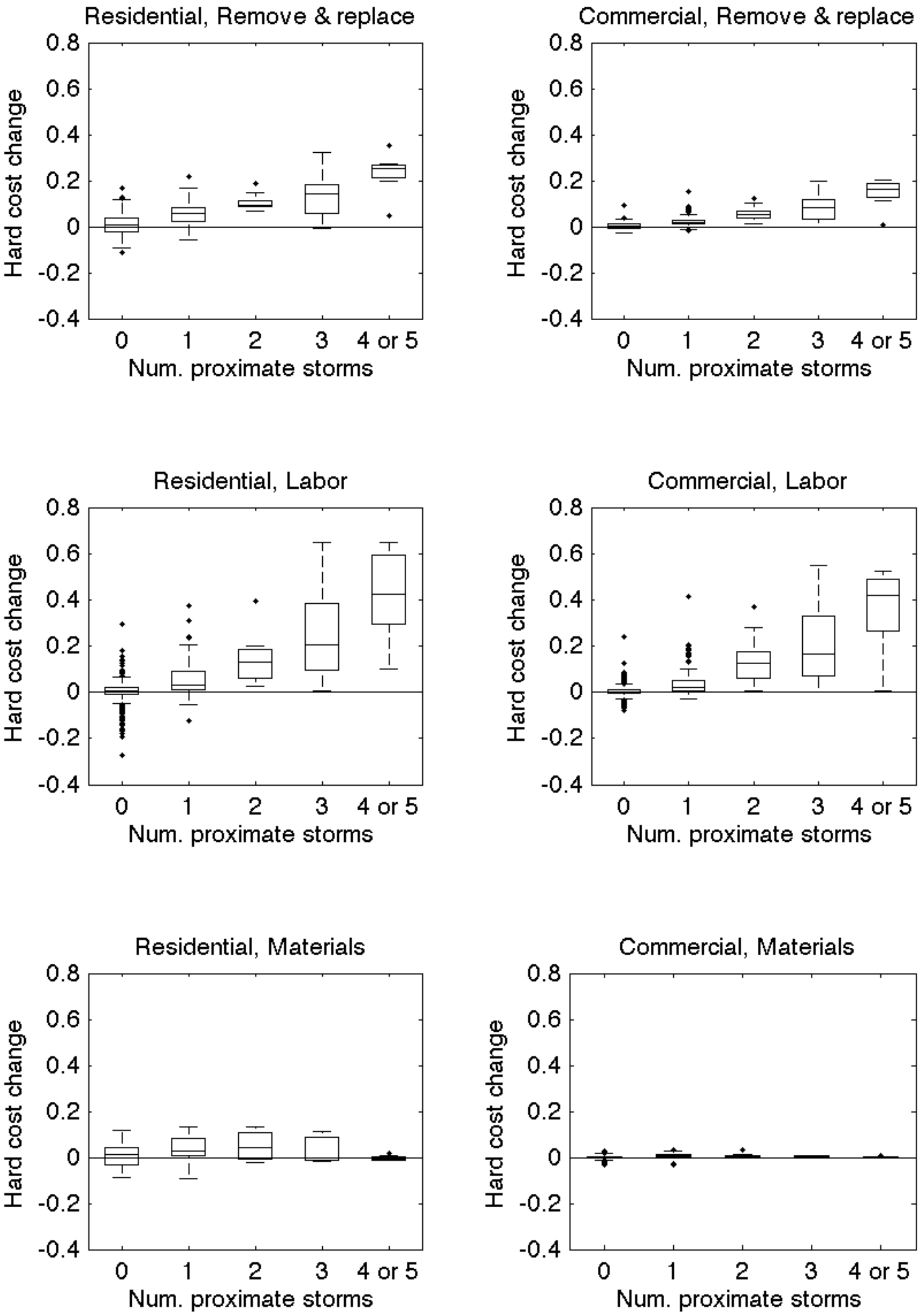

Figure 5: Observed hard cost change from July to January categorized by the number of proximate storms

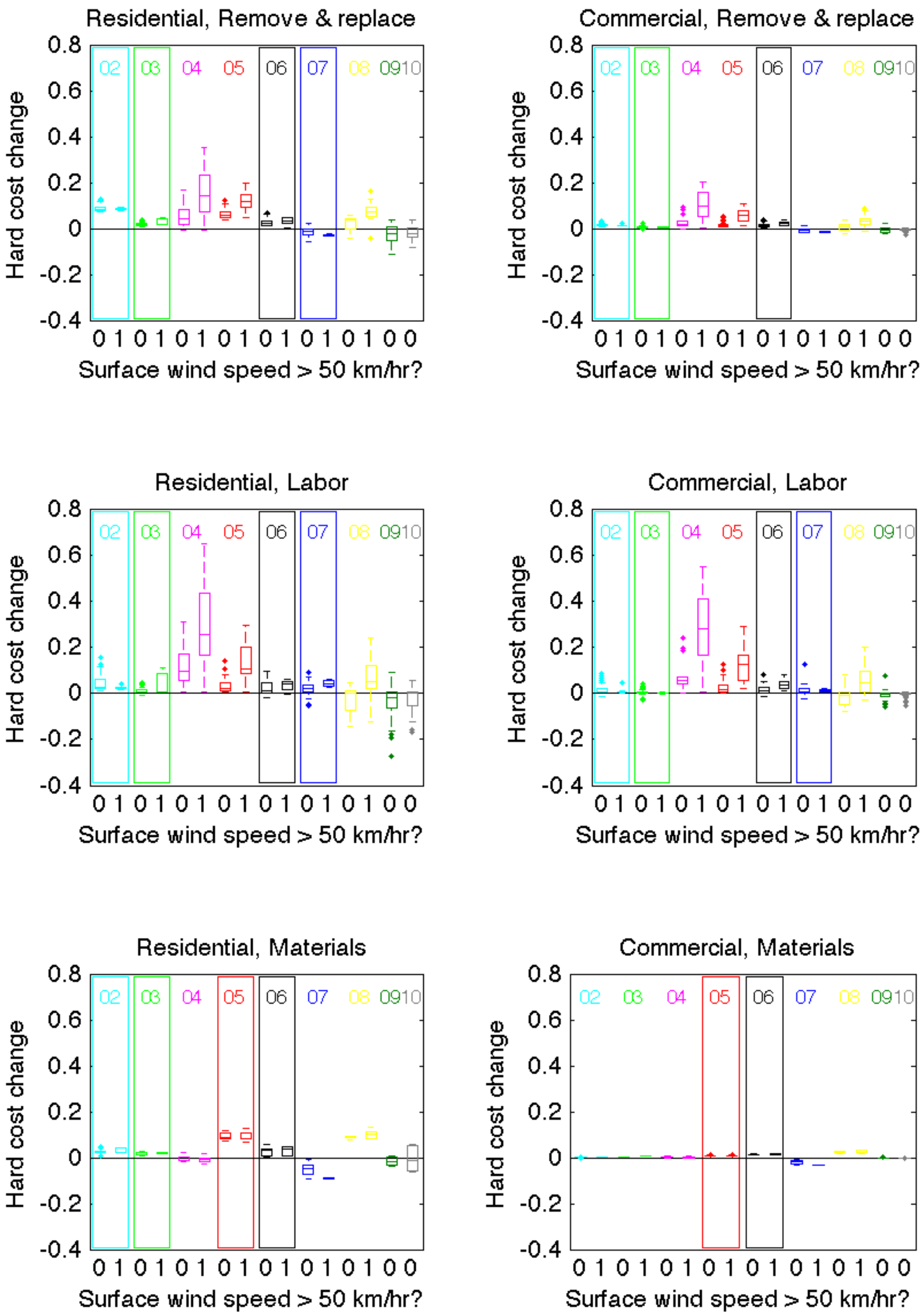

Figure 6: Observed hard cost change from July to January categorized by year. Rectangles around box-and-whiskers indicate that we could not reject the null hypothesis that, in a given year, cost changes in cities affected by a large proximate storm are drawn from the same population as cost changes in cities unaffected by such a storm.

	p-value								
Basket	*TX*	*LA*	*MS*	*AL*	*FL*	*GA*	*SC*	*NC*	*VA*
RES R&R	0.0003	0.0000	0.0002	0.0003	0.0000	0.3483	0.0392	0.1485	0.1984
COM R&R	0.0000	0.0017	0.0028	0.0001	0.0000	0.3483	0.0023	0.0026	0.9350
RES LAB	0.0001	0.0017	0.0098	0.0002	0.0000	0.7868	0.0346	0.0473	0.1984
COM LAB	0.0001	0.0371	0.0298	0.0003	0.0000	0.1314	0.0067	0.5472	0.5344
RES MAT	0.0000	0.0000	0.6693	0.1766	0.0000	0.1994	0.0236	0.0058	0.5344
COM MAT	0.0020	0.0371	0.6693	0.7096	0.0000	0.1626	0.0158	0.0001	0.1984

Table 2: P-values from hypothesis tests that observed hard cost changes in a given U.S. state do not differ between cities affected and unaffected by large proximate storms

with a similar null hypothesis that data from unaffected cities in a given state are drawn from the same underlying population as data from affected cities in that state. Tables 2 and B.2 show the p-values for all tests on the hard and market costs, respectively.

The cost changes in some states are significantly affected by a large proximate storm, while others are not. For four baskets—residential and commercial, remove & replace and labor component—states on the Atlantic Coast tend not to have significantly different cost changes for affected versus unaffected cities. In Georgia and Virginia (and, for the residential remove & replace and the commercial labor baskets, North Carolina), cost changes when there is a large proximate storm are not significantly different than cost changes when there is no such storm. This pattern does not hold for the material component baskets, however. Cost changes between affected and unaffected cities are not significantly different in Mississippi and Alabama (on the Gulf Coast) and Georgia and Virginia (on the Atlantic Coast). This observation suggests that the state, or another factor for which state is a proxy, affects the cost change in a city after a large proximate storm. In other words, there may be regional differences in how large proximate storms affect cost changes.

Finding 10 *For the remove & replace and labor component baskets, states on the Atlantic Coast tend not to have significantly different cost changes between cities in their borders affected versus unaffected by at least one large proximate storm. States on the Gulf Coast and Florida, however, tend to have significantly different cost changes between affected and unaffected cities. For the material component baskets, there is no such regional trend. The state in which a city is located may be a proxy for a factor that more directly determines cost changes regionally, such as restrictions on the movement of materials and labor or requirements regarding reconstruction.*

Figures 8 and B.5 show quarterly cost changes, as opposed to the cost changes in the second half of the year shown in all previous figures of this section. For the third and fourth quarters, we again categorize the data according to the presence of a large proximate storm in the quarter.[5] There are missing data points in Xactimate's database, especially in Florida at the beginning of the fourth quarter of 2004. To suggest what a complete data set might show, we linearly interpolate the values of missing data points and also plot the original data set augmented with the interpolated points in Figures 8 and B.5.

Cost changes in the first two quarters of the year are consistent with cost changes in the last two quarters of the year given no large proximate storm. The distributions of cost change in these quarters have small means and variances but many positive or negative outliers. These distributions are distinct from the distributions of cost change given at least one large proximate storm, which have larger medians and variances but few outliers. This provides evidence that cost changes given a large proximate storm are different than cost changes given no such storm, independent of the quarter of the year under consideration. The quarterly distributions in the second half of the year given no large proximate storm could have been distinct from the distributions in the first half of the year; in this case, it would have been difficult to associate large cost changes with large proximate storms alone, and not with some unidentified, independent economic factor of the second half of a calendar year.

Finding 11 *Cost changes given a large proximate storm are different than cost changes given no such storm, independent of the quarter of the year under consideration. In the Southeastern United States, the distributions of cost change by quarter are distinct only when there is at least one large proximate storm.*

3.2 Grouping the Data

The findings of our data exploration (Section 3.1) suggest that wind speed, number of proximate storms, U.S. state, and year are all possible explanatory variables to predict the cost change response variable. We expected variables

[5] Some tropical storms made landfall and affected cities in the United States near the end of the third quarter. The effects of these storms likely were not included in the costs reported at the beginning of the fourth quarter. Thus, we assign proximate storms from the last week of the third quarter to the fourth quarter.

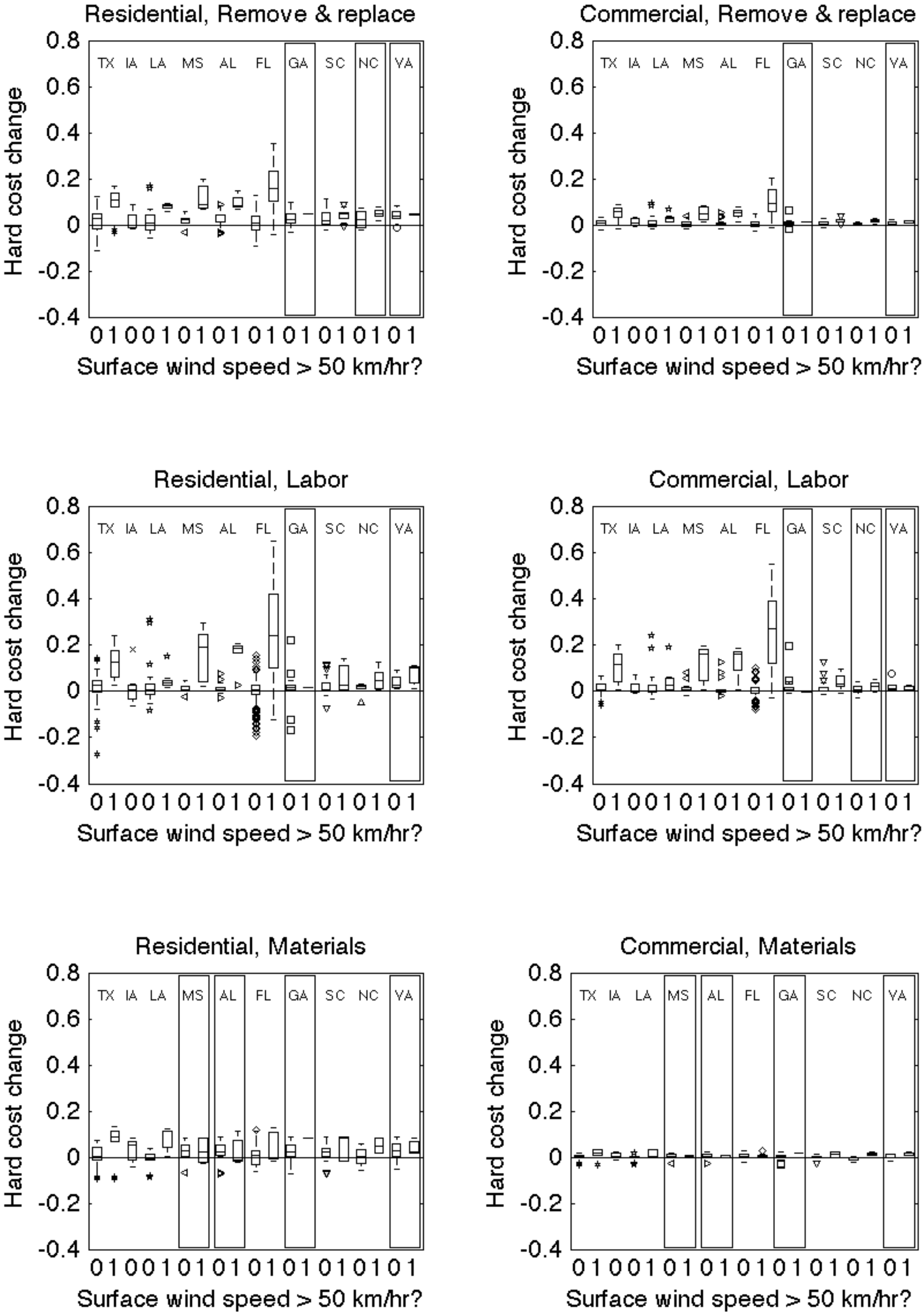

Figure 7: Hard cost change from July to January categorized by U.S. state. Rectangles around box-and-whiskers indicate that we could not reject the null hypothesis that, in a given U.S. state, cost changes in cities affected by a large proximate storm are drawn from the same population as cost changes in cities unaffected by such a storm.

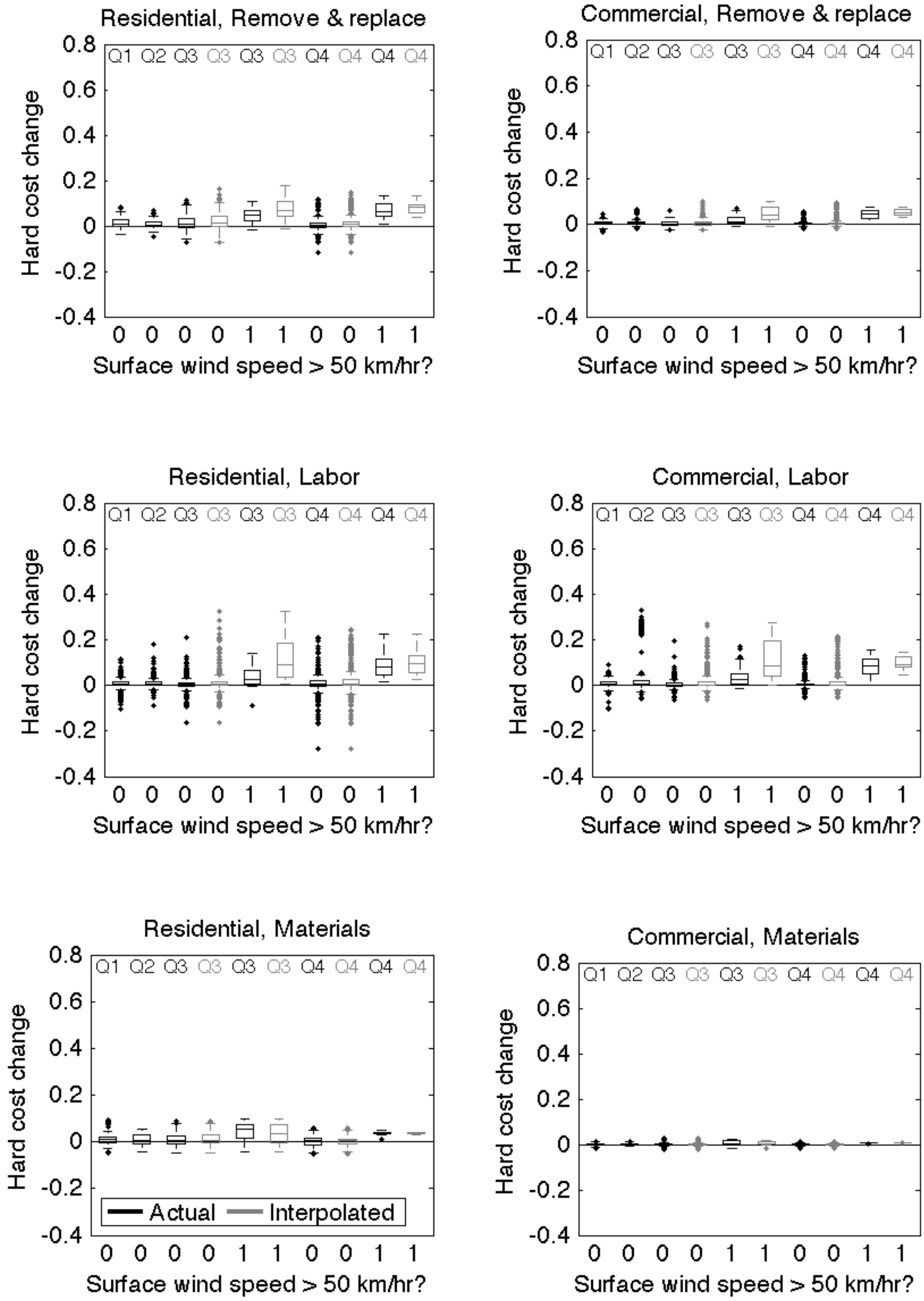

Figure 8: Observed hard cost change by quarter. When there are missing data in the Xactimate databases, we linearly interpolate the missing datum from the preceding and succeeding quarters. The data as extracted from the Xactimate databases form the "actual" data sets, and these data sets augmented with the linearly interpolated points form the "interpolated" data sets.

such as these to affect demand surge based on past, qualitative explanations of demand surge (see Section 1 or Olsen and Porter (2010)), and the data support an attempt to quantify these relationships. However, we cannot immediately proceed to a regression analysis of our response variable on the proposed explanatory variables: the statistical "unit of analysis" is different among the variables. Wind speed, number of proximate storms, and cost change are quantities measured at each city, but we expect variation from city-to-city, as well as from state-to-state and from year-to-year. We would like to determine how much of the total variation in cost change is at the city level versus the state level versus the annual level. To do this we utilize the techniques of multilevel models, also known as mixed effects models, in our regression analyses.

Generally speaking, data describe some characteristic of a unit of analysis, or an object under study. A unit of analysis may be a school student, a flower in a garden, or a city on the Gulf or Atlantic Coast of the United States. An individual can be grouped with other, similar individuals, and comparisons between groups may be of interest, in addition to comparisons among individuals. Students can be grouped by classroom or teacher, flowers grouped by species, and cities grouped by state. Groups can be grouped themselves into hierarchies of groups, or individuals can be cross-classified into several, non-hierarchical groups, until all units of analysis of interest are formed. Classrooms can be grouped hierarchically by school and then by district. Individual flowers can be grouped by species or by patch of garden, resulting in a cross-classification scheme. Economic data for cities can be grouped by state or by year, also resulting in cross-classification. Repeated grouping in this way forms multiple levels of analysis units, and variability can be studied within each level but also across levels. (See, for example, Gelman and Hill (2007) or Hox (2010) for further explanations of the theory and applications of multilevel models.)

In our data exploration, we could group the data by year, by state, or by state and year simultaneously. We immediately find, however, that some groups have little data. Instead of groups by state, we regionalize the states as the Gulf Coast (Texas, Louisiana, Mississippi, and Alabama; 17 cities), Florida (25 cities), and the Atlantic Coast (Georgia, South Carolina, North Carolina, and Virginia; 10 cities). Figures 9 and 10 show the residential remove & replace basket, hard cost change versus gradient wind speed, grouped by these regions and by year. (Data in this section include interpolated values when the costs are missing in the Xactimate databases.) Plots of the residential labor component (Figures B.6 and B.7), the commercial remove & replace (Figures B.8 and B.9), and the commercial labor component data (Figures B.10 and B.11) are similar; we use Figures 9 and 10 as representative of the data for these other baskets. Similarly, Figures 11 and 12 show the residential material component basket, hard cost change versus wind speed, grouped by region and by year. Figures B.12 and B.13 plot the commercial material component data.

Figures 9 and 10 suggest that the primary source of variability in the cost change of the remove & replace or labor component basket is the year, not the geographic region. In many years, specifically all years except 2004, 2005, and 2008, the cost changes have a relatively small variation about a mean, which is different from year-to-year, and the cost changes when there is a proximate storm don't seem different than the cost changes when there is no proximate storm. In 2004, 2005, and 2008, the cost changes when there is a proximate storm are distinct from those without a proximate storm. A linear relationship between cost change and wind speed for these data seems justified in the absence of additional information.

Finding 12 *For the remove & replace and labor component baskets, the primary source of variability in the cost change is the year, not the region of the Southeastern United States.*

The city-to-city variability within each of the three regions is somewhat smaller than the region-to-region variability. In other words, when we compare the subplots within a single row, it can be difficult to distinguish the data's scatter within one region from the scatter within another region. A clear distinction is the data from Florida in 2004. Cost changes in the Atlantic and Gulf States in 2004 do not exceed 20 percent, whereas cost changes in many Florida cities are at least 20 percent. We also consider whether data from the most northern cities in Florida are distinct from the peninsular Florida data. The unfilled markers in this set of figures represent Florida cities not on its peninsula (specifically, Jacksonville, Panama City, Pensacola, and Tallahassee). Given these data, there is no reason to distinguish peninsular versus Northern Florida cities when analyzing cost changes. Also, consider the data from 2009 (third row of Figure 10), for example. If the data were combined into a single data set, the variance of this data set would be larger than the variance of the individual data sets. These observations suggest that data within a given region and year can be modeled with a smaller residual variance than the data without assignment to such groups.

Finding 13 *For groups defined by region and year, the within-group spreads of the available data are much smaller than the spreads if the data were grouped by year alone. This adds support to the choice of grouping by region and year, even though the year-to-year variability seems to exceed the region-to-region variability.*

Like the remove & replace and labor component baskets, the primary source of variability in the cost change of material component baskets seems to be the year not the region (Figures 11 and 12). In all years, the material component cost changes have a relatively small variation about a mean, which is different from year to year, and the

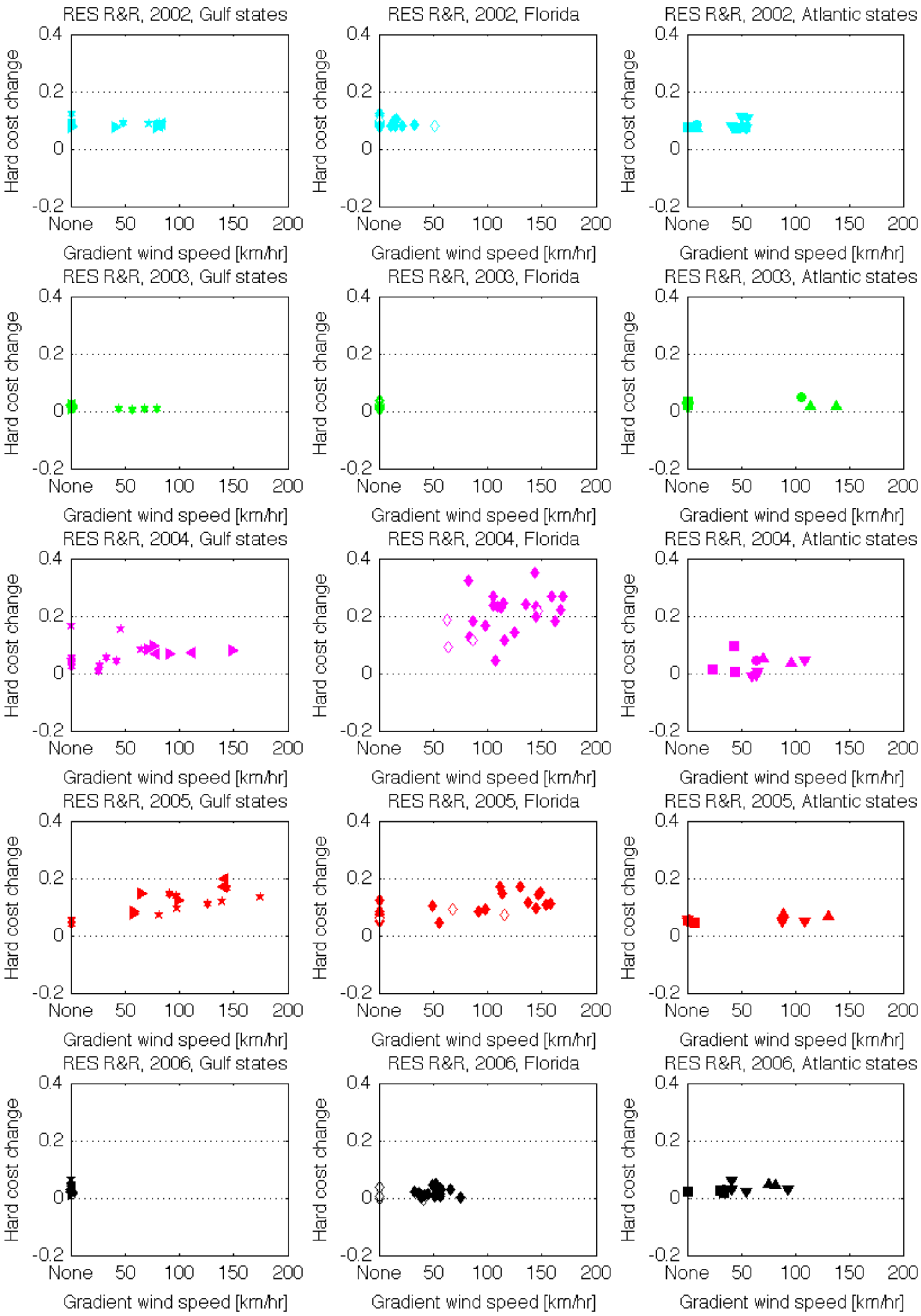

Figure 9: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the residential remove & replace basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

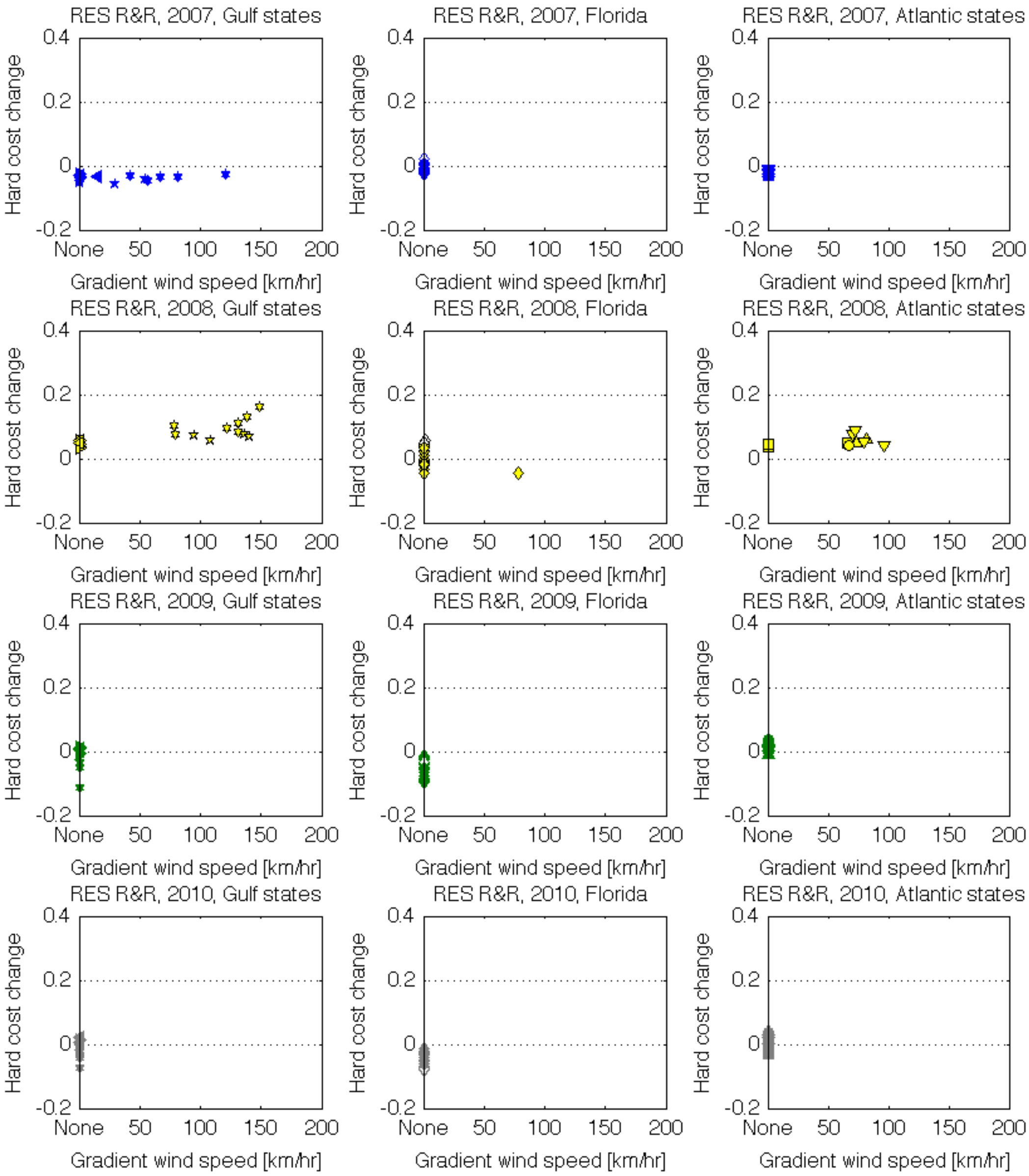

Figure 10: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the residential remove & replace basket.

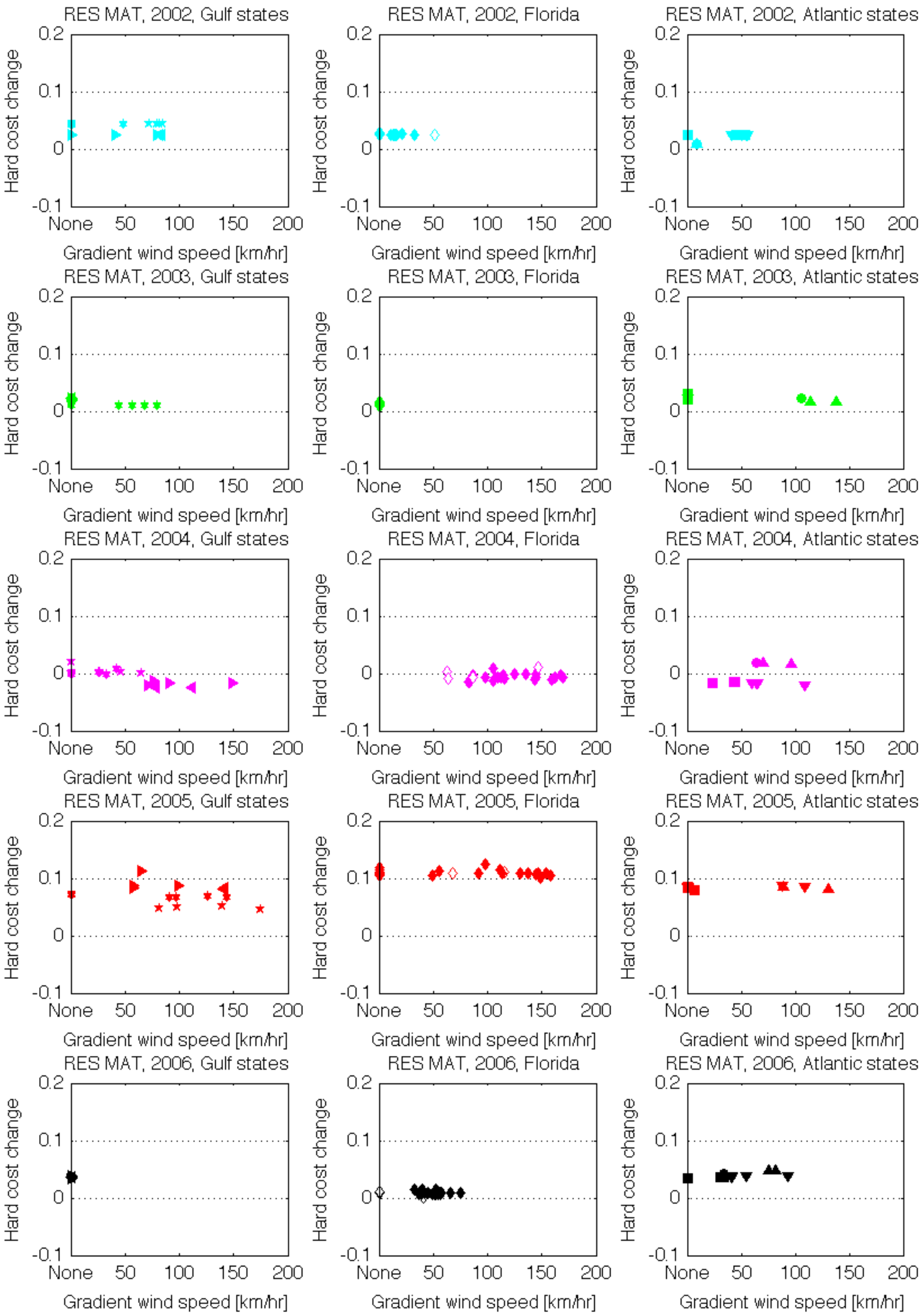

Figure 11: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the residential material component basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

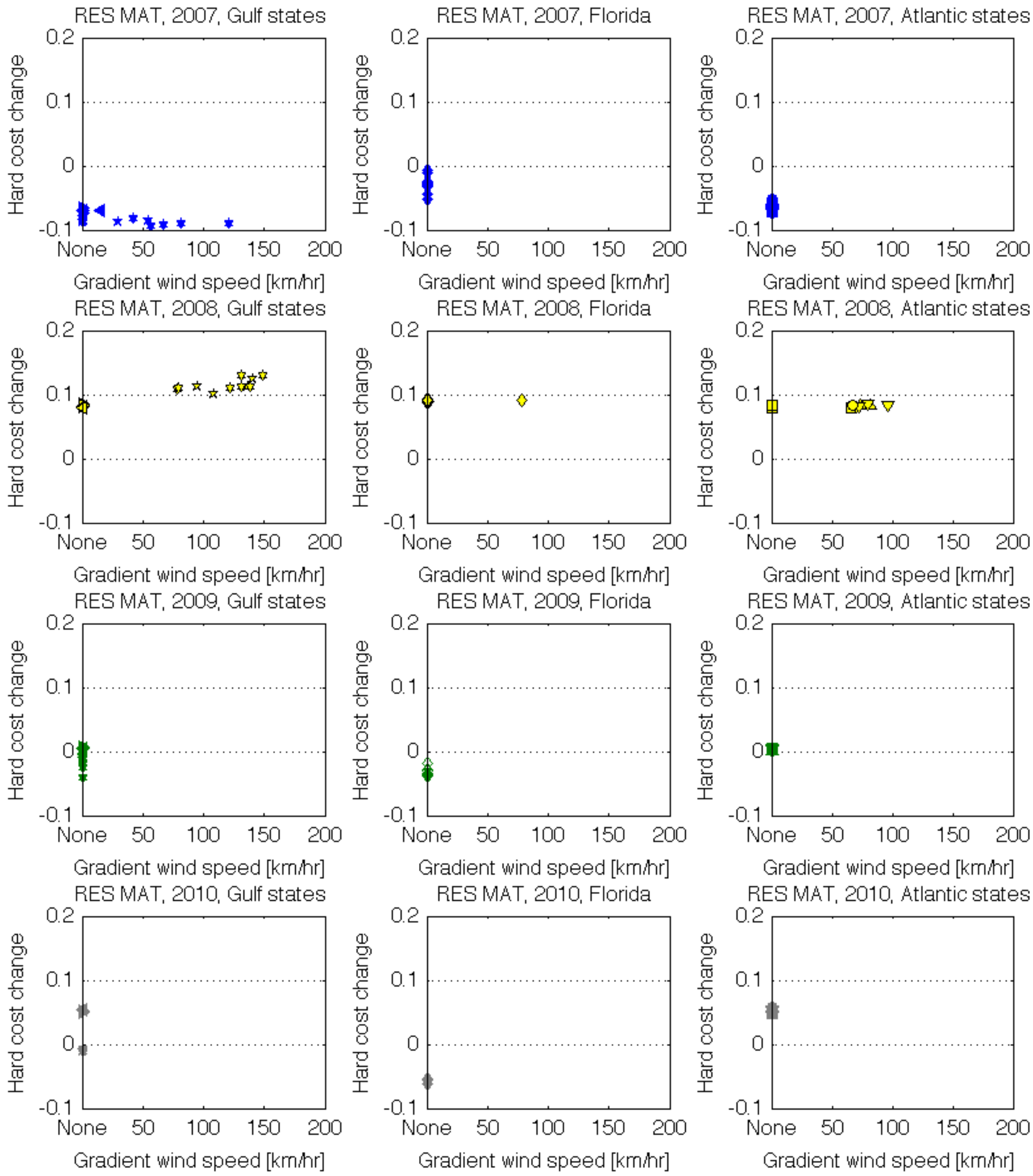

Figure 12: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the residential material component basket.

cost changes when there is a proximate storm don't seem different than the cost changes when there is no proximate storm. A distribution of cost change given the year may be sufficient to model the variability in material component cost changes, independent of the region of the United States and the wind speed.

Finding 14 *For the material component baskets, the cost changes within a year have a small variability about a mean, independent of wind speed, but the mean value can vary substantially from year to year.*

3.3 Quantitative Models

The development of our multilevel, quantitative models of cost changes follows procedures typical of any model development. In the two previous sections, we explored the data. In this section, we begin by applying a simple linear regression model and argue that it is inappropriate. Then we propose two sets of multilevel models of the data and evaluate how well each proposed model represents the available data. The first set of models assumes a constant residual variance for all groups, and the second set allows the residual variance to be different for different groups. Finally, we select the best model for each basket, given the data and our proposed models. We follow the suggestions of Pinheiro and Bates (2000, Chapter 4) for multilevel model development, and all statistical analyses are done with R (R Development Core Team, 2009).

We can use a simple linear regression equation to model the city-level data, irrespective of U.S. region or year. The data in this section include interpolated values when the costs are missing in the Xactimate databases. We assume that the data are independent and identically normally distributed with a mean that is a linear function of gradient wind speed and a constant variance. As an equation, for city i:

$$\Delta_i = \alpha + \beta w_i + \epsilon_i, \tag{1}$$

where Δ_i is the cost change from July to the following January, α and β are model parameters, w_i is the gradient wind speed, and $\epsilon_i \sim \mathcal{N}\left(0, \sigma^2\right)$, read "$\epsilon_i$ is normally distributed with mean zero and variance σ^2."

Figures 13 and 14 show the residuals of Equation 1 with parameter values estimated for the six baskets. In the subplots at left, the residuals do not appear to be symmetric about zero, and especially in the labor and material component baskets, many residuals form clear lines. In the normal probability graphs at right, the residuals clearly deviate from what we would expect if they were drawn from a normal distribution. Thus, the residuals violate the assumptions of the simple linear regression model.

More importantly, these plots do not show the flaw in applying a standard regression to grouped data. When we use standard regression on data that should be grouped, the standard errors of the estimated parameter values are smaller than they would be with multilevel regression. (See Goldstein (2003, Section 2.7) or Hox (2010, Section 1.2).) These standard errors are used to test the significance of the parameters in the model. Since the standard error calculated by standard regression is smaller than the standard error from multilevel regression, a parameter may be deemed "significant" in the model, when in fact, the data do not support such a finding. This (fictitious) significance results from the estimation algorithm, not from the data.

We now apply Equation 1 to the data in each region, to the data in each year, and to the data in each combination of region and year. We do this to suggest how the intercepts and slopes vary by grouping, not to find an appropriate model of the data. For example, Figures 15 and B.14 through B.18 show Equation 1 applied to the data in each year. Note that the estimated intercepts and slopes vary by year. The intercept for the 2002, residential remove & replace data is larger than the intercept in 2003, and the slope in 2004 is larger than the slope in any other year. The remove & replace and labor component data appear to be independently distributed about the simple regression line, however, the material component data do not. Figures B.15 and B.18 show systematic deviations from the simple regression line by U.S. state. These deviations are most obvious in the 2005 residential material component data. Cost changes in Florida (diamond-shaped symbols) are consistently larger than the values predicted from the regression using all 2005 data. Cost changes in Texas (six-pointed stars) and in Louisiana (five-pointed stars) are each consistently smaller than the predicted values.

We can evaluate how the estimated intercepts and slopes vary across different data groupings by comparing their confidence intervals across groups. Figures 16 and B.19 through B.23 show confidence intervals for the intercepts and slopes of Equation 1 applied to data grouped by region, by year, and by region and year simultaneously. No single value for the intercept (or single value for the slope) is inside the confidence intervals of all groups, whether the data are grouped by region, year, or both. This suggests that both the intercept and slope are stochastic themselves, not parameters with single values, and the values of the intercept and slope vary by group.

Finding 15 *Assuming a linear relationship between wind speed and cost change, no single value of the intercept, nor single value of the slope, is appropriate for all groups of data. The intercept and slope vary across groups, and lacking additional, group-level explanatory variables, we assume the intercept and slope are stochastic.*

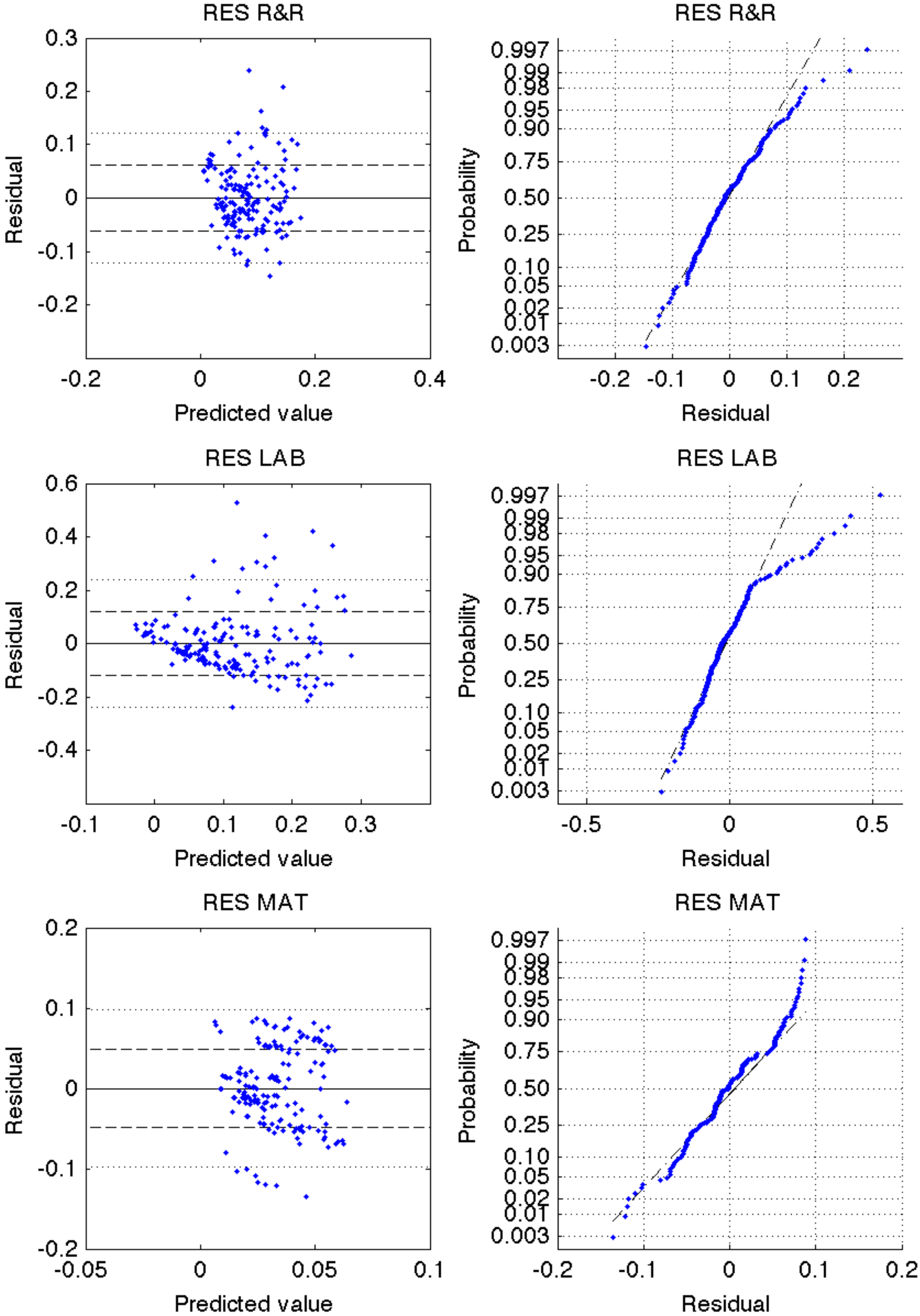

Figure 13: Residuals of hard cost change from July to January versus wind speed from a simple linear regression for the residential baskets

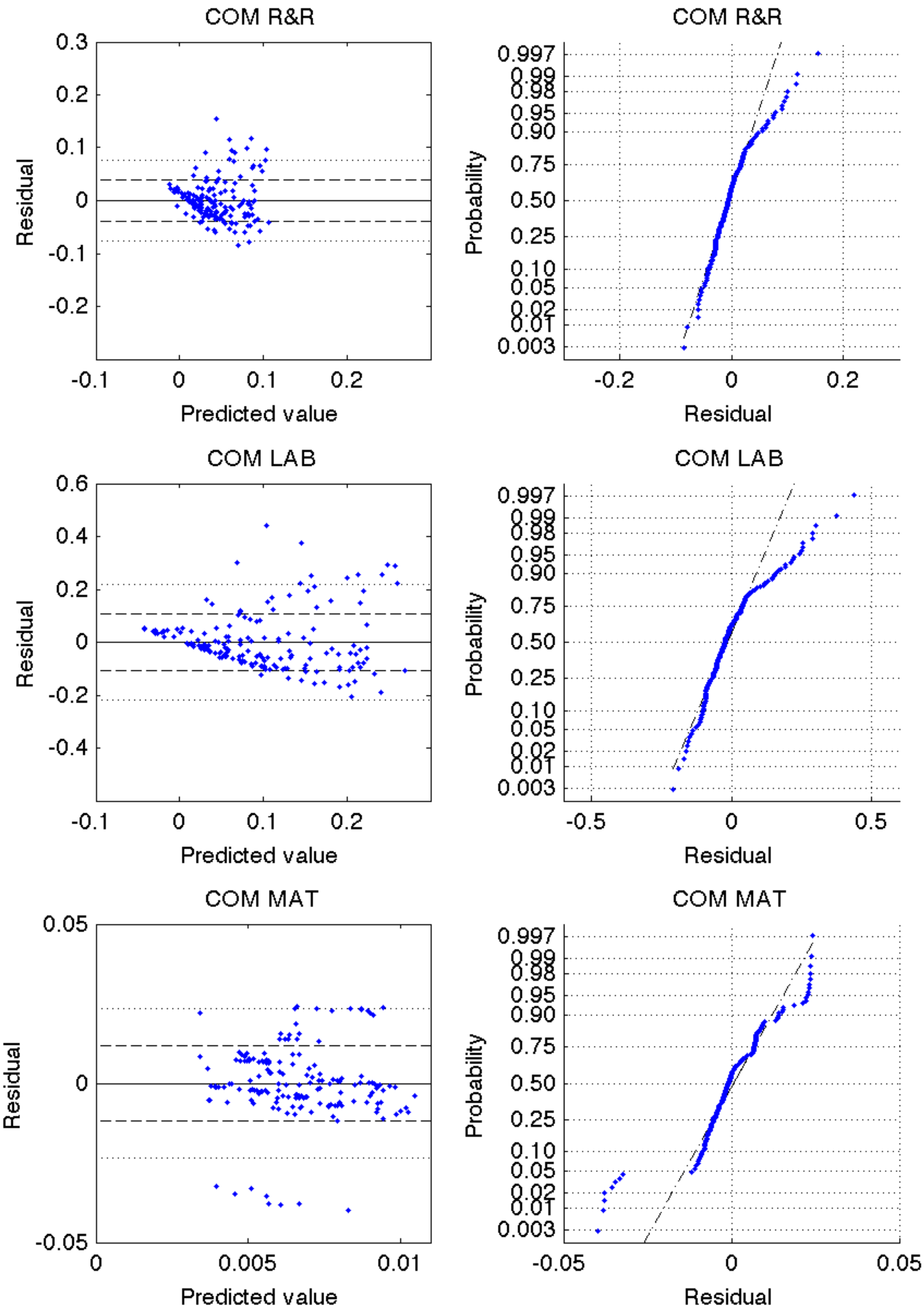

Figure 14: Residuals of hard cost change from July to January versus wind speed from a simple linear regression for the commercial baskets

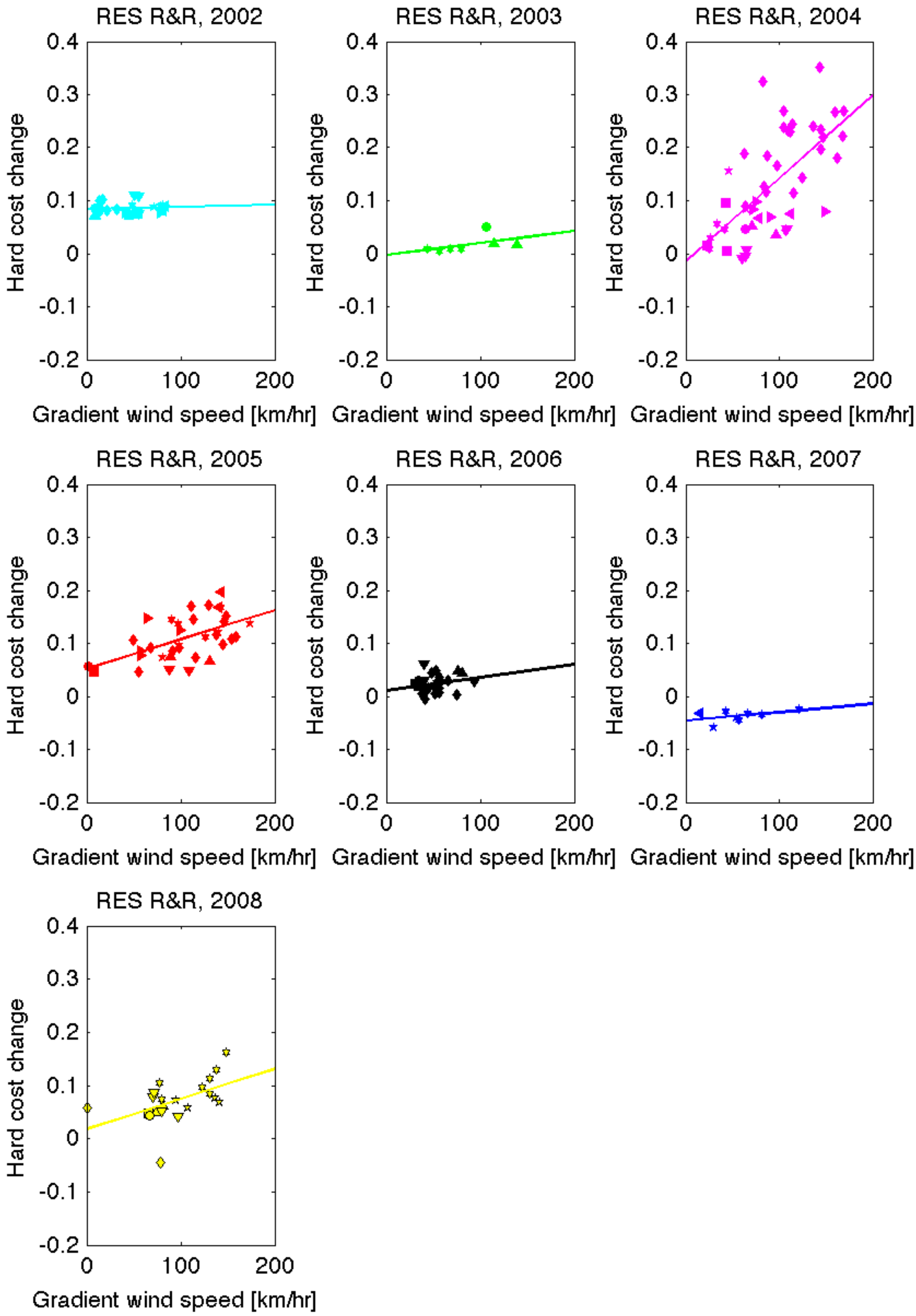

Figure 15: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the residential remove & replace basket

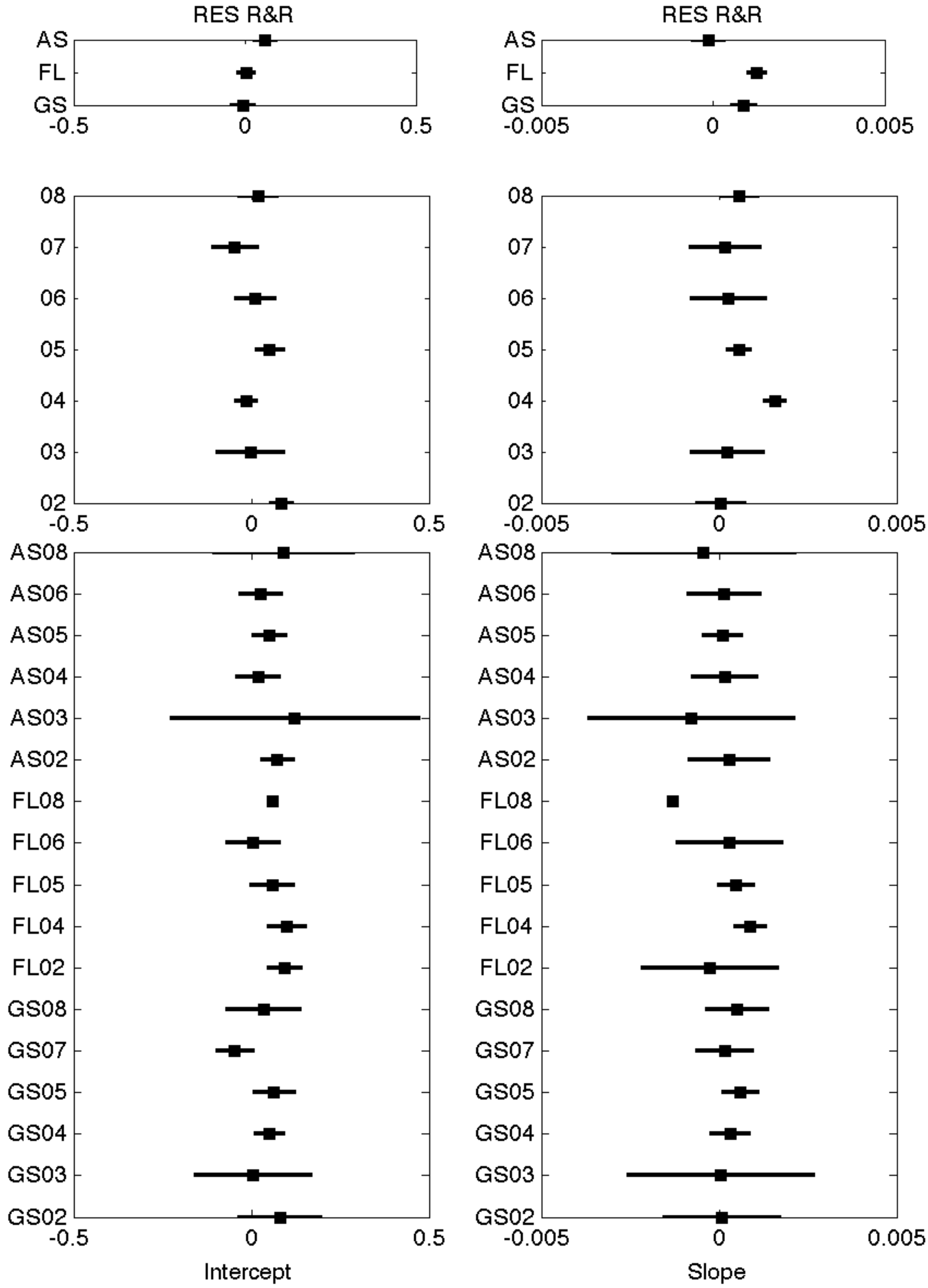

Figure 16: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the residential remove & replace basket. "AS" is Atlantic States; "FL" is Florida; and "GS" is Gulf States.

3.3.1 Proposed Models Assuming a Constant Variance

We now propose a series of multilevel models that formalize our suspicions about the structure of the cost change data. In particular, at the city level, we suspect that the wind speed of a proximate storm, the number of proximate storms, and cost changes in the first half of the year may predict the cost change in the last half of the year. Formally, we propose a general model structure to account for these suspicions:

$$\Delta_i \quad = \quad \alpha_{jk} + \beta_{jk} w_i + \gamma_{jk} n_{ci} + \delta_{jk} w_i n_{ci} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_i, \tag{2}$$

where

- in year j and region k,

$$\begin{bmatrix} \alpha_{jk} \\ \beta_{jk} \\ \gamma_{jk} \\ \delta_{jk} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_\alpha \\ \mu_\beta \\ \mu_\gamma \\ \mu_\delta \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 & \rho_{\alpha\beta}\sigma_\alpha\sigma_\beta & \rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma & \rho_{\alpha\delta}\sigma_\alpha\sigma_\delta \\ \rho_{\alpha\beta}\sigma_\alpha\sigma_\beta & \sigma_\beta^2 & \rho_{\beta\gamma}\sigma_\beta\sigma_\gamma & \rho_{\beta\delta}\sigma_\beta\sigma_\delta \\ \rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma & \rho_{\beta\gamma}\sigma_\beta\sigma_\gamma & \sigma_\gamma^2 & \rho_{\gamma\delta}\sigma_\gamma\sigma_\delta \\ \rho_{\alpha\delta}\sigma_\alpha\sigma_\delta & \rho_{\beta\delta}\sigma_\beta\sigma_\delta & \rho_{\gamma\delta}\sigma_\gamma\sigma_\delta & \sigma_\delta^2 \end{bmatrix} \right),$$

  and the covariance matrix represents the variation of intercept and slope coefficients between groups defined by year and region;

- $\epsilon_i \sim \mathcal{N}\left(0,\ \sigma^2\right)$ and represents the residual, or unexplained, variation;

- at city i, Δ_i is the basket cost change from July to the following January; w_i is the gradient wind speed in km/hr; n_{ci} is the number of proximate storms; Δ_{1i} and Δ_{2i} are the basket cost changes in the first and second quarters, respectively;

- and μ_α, μ_β, μ_γ, μ_δ, σ_α^2, σ_β^2, σ_γ^2, σ_δ^2, $\rho_{\alpha\beta}$, $\rho_{\alpha\gamma}$, $\rho_{\alpha\delta}$, $\rho_{\beta\gamma}$, $\rho_{\beta\delta}$, $\rho_{\gamma\delta}$, ϕ_1, ϕ_2, ϕ_3, and σ^2 are model parameters.

The subscript jk indicates that the variable is defined for the grouped data from year j and region k.

With its estimated parameter values, the full model of cost change (Equation 2), or any reduced model formed by setting some parameter values equal to zero, is a proposed model. For any set of proposed models, we would like to select the model that best represents our available data. There are many techniques for model selection, but all are based on the premise that a model should be only as complex as justified by the observed data. For each proposed model, we calculate Akaike's "an information criterion" (AIC). This criterion compares the model complexity (measured by the number of parameters) against the likelihood of observing the available data given the proposed model. As an equation:

$$\text{AIC} = 2K - 2\log_e\left(\mathcal{L}\right), \tag{3}$$

where K is the number of model parameters and $\mathcal{L}$ is the likelihood function of the proposed model evaluated at the likelihood's maximum point. (The details of constructing a likelihood function are not important here. See, for example, Burnham and Anderson (2002, Chapters 1–2), for a complete discussion of the likelihood function and model selection techniques.) Calculating AIC for each proposed model allows us to rigorously compare its complexity to the likelihood of observing the available data. Ideally, a proposed model should have few parameters and a large likelihood, and thus the "best" model, among those proposed, has the smallest AIC value.

We use an iterative process to generate proposed models for the cost change of each of the six baskets of repair costs. We first apply Equation 2 to the data, finding the Student's t-value for each coefficient and the AIC value for the model.[6] We then remove the explanatory variable whose coefficient has the smallest absolute t-value, and we calculate the new Student's t-values and AIC value for the reduced model. We continue this process until we find the combination of explanatory variables that has the smallest AIC value. Then, we consider models where the intercept and coefficients are not stochastic, if these parameters remained in the reduced model. For example, we set $\alpha_{jk} = \alpha$, a deterministic value, and thus $\sigma_\alpha^2 = 0$. We try different combinations of deterministic and stochastic coefficients until we find one that minimizes AIC among those reduced from the full model in Equation 2. A model we have not yet considered may have an even smaller AIC value.

Tables 3 and B.3 through B.7 show our proposed models and the log-likelihood (that is, $\log_e\left(\mathcal{L}\right)$ in Equation 3) and AIC value for each model. Following Burnham and Anderson (2002), we also report the difference between each

[6] When calculating AIC values, we use full, not restricted, maximum likelihood estimation. Full maximum likelihood estimation allows us to compare models that differ in the deterministic part as well as the stochastic part (Hox, 2010, Section 3.2.3). After we select the best combination of explanatory variables—that is, the best reduced model—we use restricted maximum likelihood to estimate parameter values and standard errors of the estimates. The restricted maximum likelihood method produces estimates with less bias than those estimated using the full maximum likelihood method (Hox, 2010, Section 3.1.1).

model's AIC value and the minimum AIC value over the proposed models in each table, Δ_{AIC}, and we calculate the weight for each model as:

$$\text{weight} = \frac{\exp\left(-\frac{1}{2}\Delta_{\text{AIC}}\right)}{\sum_{m=1}^{M} -\frac{1}{2}\Delta_{\text{AIC}_m}},$$

where M is the number of proposed models for a basket of repair costs. In the tables, the proposed models are ordered by increasing AIC values. Burnham and Anderson (2002, Section 2.6) suggest that a model with Δ_{AIC} roughly between 0 and 2 has "substantial" empirical support. Using this guideline, we find that several proposed models for each basket of repair costs can have substantial support. We consider the proposed models for each basket, and report the model with the smallest AIC value, in the following paragraphs.

City-level model						Log-likelihood	AIC	Δ_{AIC}	weight
α_{jk}			$+\delta w_i n_{ci}$		$+\epsilon_i$	316.1	-624.3	0.0	0.4189
α_{jk}			$+\delta_{jk} w_i n_{ci}$		$+\epsilon_i$	317.7	-623.4	0.9	0.2671
α_{jk}			$+\delta_{jk} w_i n_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_i$	318.5	-623.0	1.3	0.2187
α_{jk}			$+\delta_{jk} w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_i$	318.7	-621.3	3.0	0.0935
α_{jk}	$+\beta_{jk} w_i$		$+\delta_{jk} w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_i$	318.8	-613.5	10.8	0.0019
α_{jk}	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	319.1	-602.1	22.2	0.0000
α			$+\delta_{jk} w_i n_{ci}$		$+\epsilon_i$	298.1	-588.3	36.0	0.0000

Table 3: First set of proposed models and their AIC values for hard cost change from July to January of the residential remove & replace basket

For the residential remove & replace basket, three of the seven proposed models have substantial support. These three models include: a stochastic intercept; wind speed (w) crossed with the number of proximate storms (n_c) as an explanatory variable but not either of the two variables alone; and the third model also includes the cost change in the second quarter. The two models with the smallest AIC values differ only in the coefficient of the cross term. In the first model, this coefficient is deterministic, while in the second model, the coefficient is stochastic. The model with the smallest AIC value is:

$$\Delta_i^{\text{RES R\&R}} = \alpha_{jk} + \delta w_i n_{ci} + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\mu_\alpha} = 0.04262$ (0.01161), $\sigma_\alpha^2 = 0.001946$;
- $\widehat{\delta} = 0.0001888$ (0.00002867);
- and $\sigma^2 = 0.001115$.

Throughout this report, we indicate parameter values estimated from the available data with a "hat."

For the residential labor component, three of the six proposed models have substantial support. These three models include only wind speed and the number of proximate storms as separate explanatory variables, and the first two models also include the cost change in the second quarter. The two models with the smallest AIC values have a stochastic coefficient for the number of proximate storms and differ in the coefficient of the wind speed variable. This coefficient is deterministic in the first model and stochastic in the second model. The model with the smallest AIC value is:

$$\Delta_i^{\text{RES LAB}} = \beta w_i + \gamma_{jk} n_{ci} + \phi_2 \Delta_{2i} + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\beta} = 0.0007093$ (0.0001692);
- $\widehat{\mu_\gamma} = 0.006507$ (0.01401), $\sigma_\gamma^2 = 0.001473$;
- $\widehat{\phi_2} = 0.4836$ (0.2919);
- and $\sigma^2 = 0.005112$.

For the residential material component, two of the seventeen proposed models have substantial support. These two models are reduced from Equation 2 by removing the term for cost change in the first quarter. The first model has a deterministic coefficient for the cross term of wind speed and the number of proximate storms, while the second model has a deterministic coefficient for the number of proximate storms. The model with the smallest AIC value is:

$$\Delta_i^{\text{RES MAT}} = \alpha_{jk} + \beta_{jk} w_i + \gamma_{jk} n_{ci} + \delta w_i n_{ci} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\mu_\alpha} = 0.02909$ (0.01450), $\sigma_\alpha^2 = 0.003172$;
- $\widehat{\mu_\beta} = 0.00009831$ (0.00006171), $\sigma_\beta^2 = 0.00000002228$;
- $\widehat{\mu_\gamma} = 0.008341$ (0.003595), $\sigma_\gamma^2 = 0.00003217$;
- $\widehat{\rho_{\alpha\beta}} = -0.333$, $\widehat{\rho_{\alpha\gamma}} = 0.649$, $\widehat{\rho_{\beta\gamma}} = 0.257$;
- $\widehat{\delta} = -0.00004150$ (0.00003015);
- $\widehat{\phi_2} = -0.6811$ (0.1242), $\widehat{\phi_3} = -9.960$ (4.972);
- and $\sigma^2 = 0.00004536$.

For the commercial remove & replace basket, three of the twelve proposed models have substantial support. The three models only include terms for the intercept, wind speed, and the number of proximate storms. They differ in whether the coefficients of these terms are deterministic or stochastic. In the first model, the intercept is deterministic while the two slopes are stochastic. In the second model, the intercept and the number of proximate storms coefficient are deterministic, and the wind speed coefficient is stochastic. In the third model, the number of proximate storms coefficient is deterministic, and the intercept and wind speed coefficient are stochastic. The model with the smallest AIC value is:

$$\Delta_i^{\text{COM R\&R}} = \alpha + \beta_{jk} w_i + \gamma_{jk} n_{ci} + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\alpha} = 0.01075$ (0.005553);
- $\widehat{\mu_\beta} = 0.0001377$ (0.00007725), $\sigma^2_\beta = 0.00000004489$;
- $\widehat{\mu_\gamma} = 0.001830$ (0.003541), $\sigma^2_\gamma = 0.00002657$;
- $\widehat{\rho_{\beta\gamma}} = 1.000$;
- and $\sigma^2 = 0.0003482$.

For the commercial labor component, three of the nine proposed models have substantial support. The three models all have intercept and wind speed terms, but they differ by including (or not) terms for cost changes in the first half of the year. The first model does not include cost changes from the first half of the year. The second model includes the cost change in the second quarter, and the third model includes a cross term of cost changes in the first and second quarters. The coefficients in the first model are stochastic. The model with the smallest AIC value is:

$$\Delta_i^{\text{COM LAB}} = \alpha_{jk} + \beta_{jk} w_i + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\mu_\alpha} = 0.02409$ (0.01398), $\sigma^2_\alpha = 0.0009932$;
- $\widehat{\mu_\beta} = 0.0004050$ (0.0002065), $\sigma^2_\beta = 0.0000002853$;
- $\widehat{\rho_{\alpha\beta}} = 1.000$;
- and $\sigma^2 = 0.002835$.

For the commercial material component, three of the seventeen proposed models have substantial support. The third model is reduced from the full model in Equation 2 by only removing the variable for the number of proximate storms. In the first model, the coefficient of the cross term for wind speed and the number of proximate storms is deterministic while all other stochastic coefficients in the full model remain so. The second model is the unreduced, full model. The model with the smallest AIC value is:

$$\Delta_i^{\text{COM MAT}} = \alpha_{jk} + \beta_{jk} w_i + \gamma_{jk} n_{ci} + \delta w_i n_{ci} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_i,$$

with the following estimates of parameter values and standard errors in parentheses:

- $\widehat{\mu_\alpha} = 0.005629$ (0.003143), $\sigma^2_\alpha = 0.0001504$;
- $\widehat{\mu_\beta} = 0.000004918$ (0.00001385), $\sigma^2_\beta = 0.000000001476$;
- $\widehat{\mu_\gamma} = 0.0002304$ (0.0007453), $\sigma^2_\gamma = 0.000001688$;
- $\widehat{\rho_{\alpha\beta}} = -0.035$, $\widehat{\rho_{\alpha\gamma}} = 0.291$, $\widehat{\rho_{\beta\gamma}} = 0.714$;
- $\widehat{\delta} = -0.000001770$ (0.000006289);
- $\widehat{\phi_1} = -0.1326$ (0.1097), $\widehat{\phi_2} = -0.3804$ (0.1771), $\widehat{\phi_3} = 92.60$ (39.01);
- and $\sigma^2 = 0.000001684$.

We now evaluate the model with the smallest AIC value for each of the six baskets of repair costs. Figures 17 and 18 show residual plots of these models applied to their corresponding data sets. The residuals of the models are not symmetric about zero. The remove & replace and labor component models tend to under-predict the cost change. The residuals of labor and material component models are clustered in the plots of residual versus predicted value, and they clearly deviate from a normal distribution when plotted on a normal probability scale. These residuals look worse than the residuals from applying a simple linear regression model (compare with Figures 13 and 14). However, the data may not violate the assumptions of the multilevel regression, as they do in the simple linear regression.

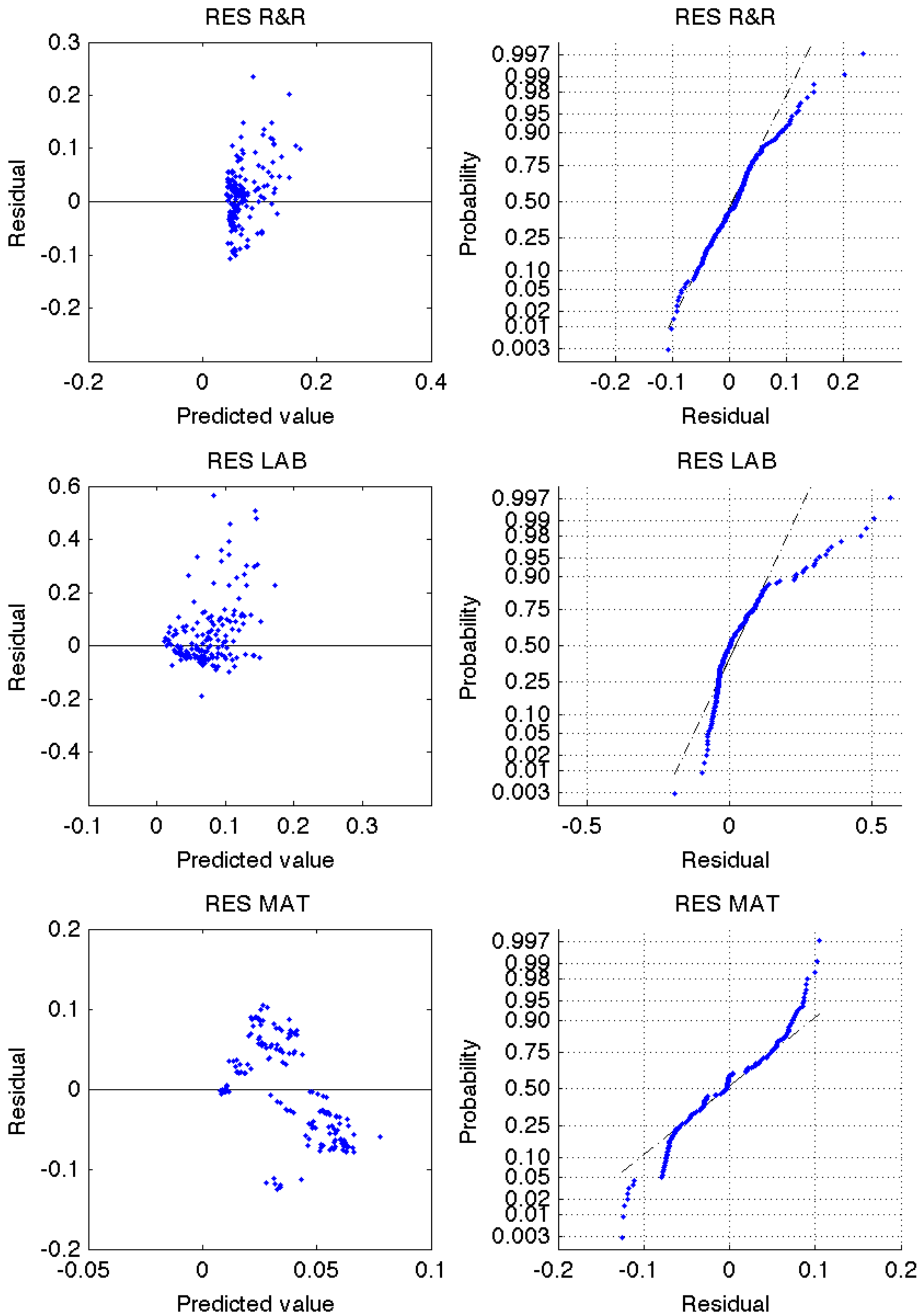

Figure 17: Residuals of hard cost change from July to January versus wind speed from the model with the smallest AIC value in the first set of multilevel regression models for the residential baskets

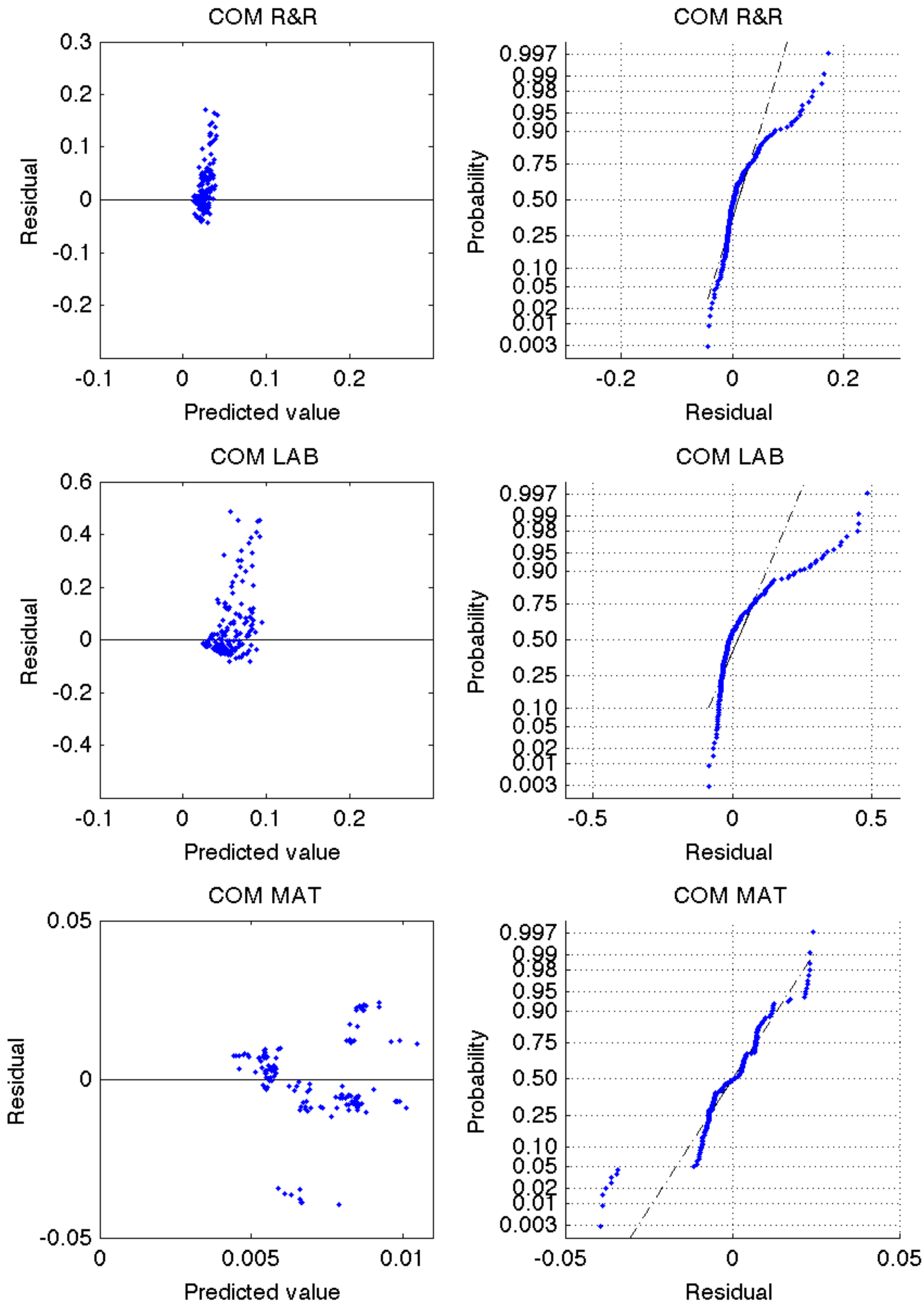

Figure 18: Residuals of hard cost change from July to January versus wind speed from the model with the smallest AIC value in the first set of multilevel regression models for the commercial baskets

3.3.2 Proposed Models Assuming a Group-Dependent Variance

The multilevel regression in Equation 2 assumes that, within a group (defined as cities in year j and region k), the data are independently and identically normally distributed about the mean value, and the within-group variance, σ, is constant and has the same value for all groups. However, Figures 15 and B.14 through B.18 suggest that σ varies between groups. We can allow this variation by proposing a second general form for the cost change of a basket of repair costs:

$$\Delta_i \quad = \quad \alpha_{jk} + \beta_{jk} w_i + \gamma_{jk} n_{ci} + \delta_{jk} w_i n_{ci} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_{jk}, \tag{4}$$

where

- in year j and region k, $\alpha_{jk} \sim \mathcal{N}\left(\mu_\alpha,\ \sigma_\alpha^2\right)$, $\beta_{jk} \sim \mathcal{N}\left(\mu_\beta,\ \sigma_\beta^2\right)$, $\gamma_{jk} \sim \mathcal{N}\left(\mu_\gamma,\ \sigma_\gamma^2\right)$, $\delta_{jk} \sim \mathcal{N}\left(\mu_\delta,\ \sigma_\alpha^2\right)$, and α_{jk}, β_{jk}, γ_{jk}, and δ_{jk} are independently distributed;
- $\epsilon_{jk} \sim \mathcal{N}\left(0,\ \sigma_{jk}^2\right)$ and represents the residual, or unexplained, variation within group jk;
- μ_α, μ_β, μ_γ, μ_δ, σ_α^2, σ_β^2, σ_γ^2, σ_δ^2, ϕ_1, ϕ_2, ϕ_3, and σ_{jk}^2 are model parameters;
- and all other symbols are defined after Equation 2.

Note that the only differences between Equations 2 and 4 are the distributions of the stochastic coefficients and different residual variances for different groups. We assume the stochastic coefficients are independently distributed because, when we allowed them to be jointly distributed, the solutions did not converge.

We again propose a series of models, now based on Equation 4. Tables 4 and B.8 through B.12 show the proposed models, as well as their log-likelihoods, AIC values, differences between each model's AIC and the minimum AIC, and model weights. As before, the models with the smallest AIC values are at the top of each table, and for some baskets of repair costs, more than one model has "substantial empirical support." The AIC values in the second sets of proposed models are all smaller than the smallest AIC values in the corresponding first sets of models.

City-level model						Log-likelihood	AIC	Δ_{AIC}	weight
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1 \Delta_{1i}$	$+\epsilon_{jk}$	396.3	-746.6	0.0	0.9882
α	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1 \Delta_{1i}$	$+\epsilon_{jk}$	391.9	-737.7	8.9	0.0117
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i}$	$+\epsilon_{jk}$	387.0	-727.9	18.7	0.0001
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_{jk}$	385.4	-724.8	21.8	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$		$+\epsilon_{jk}$	382.1	-720.1	26.5	0.0000
α_{jk}	$+\beta w_i$		$+\delta w_i n_{ci}$		$+\epsilon_{jk}$	380.4	-718.9	27.7	0.0000
α	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1 \Delta_{1i}$	$+\epsilon_{jk}$	360.0	-673.9	72.7	0.0000
α	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1 \Delta_{1i}$	$+\epsilon_{jk}$	355.9	-665.9	80.7	0.0000

Table 4: Second set of proposed models and their AIC values for hard cost change from July to January of the residential remove & replace basket

We deem the best models to be those with the smallest AIC values among the two sets of proposed models. We have no reason to select a model with substantial support but not with the smallest AIC value. The best models are:

$$\Delta_i^{\text{RES R\&R}} = \alpha_{jk} + \beta w_i + \gamma n_{ci} + \delta w_i n_{ci} + \phi_1 \Delta_{1i} + \epsilon_{jk}; \tag{5}$$

$$\Delta_i^{\text{RES LAB}} = \alpha + \beta_{jk} w_i + \gamma n_{ci} + \delta w_i n_{ci} + \phi_2 \Delta_{2i} + \epsilon_{jk}; \tag{6}$$

$$\Delta_i^{\text{RES MAT}} = \alpha_{jk} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_{jk}; \tag{7}$$

$$\Delta_i^{\text{COM R\&R}} = \alpha_{jk} + \beta w_i + \gamma n_{ci} + \delta w_i n_{ci} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_{jk}; \tag{8}$$

$$\Delta_i^{\text{COM LAB}} = \alpha_{jk} + \gamma n_{ci} + \delta w_i n_{ci} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \epsilon_{jk}; \text{ and} \tag{9}$$

$$\Delta_i^{\text{COM MAT}} = \alpha_{jk} + \phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i} + \epsilon_{jk}. \tag{10}$$

Table 5 shows the estimated parameter values (except σ_{jk}) for the best models.

Finding 16 *The best models, among those proposed, have a group-dependent variance. Although the models are more complex because each group has a different variance, the additional complexity is justified by the available data.*

Finding 17 *For the best models of the remove & replace and the labor component cost changes, wind speed, number of proximate storms, and a cost change in the first half of the year are all significant explanatory variables. The intercept is stochastic for the residential remove & replace, commercial remove & replace, and the commercial labor component models, while the slope is stochastic for the residential labor component model.*

Finding 18 *For the best models of the material component cost changes, wind speed and the number of proximate storms are* not *significant explanatory variables, whereas cost changes in the first half of the year are significant explanatory variables. The intercepts of these models are stochastic across the groupings by region and year.*

	RES R&R	*RES LAB*	*RES MAT*	*COM R&R*	*COM LAB*	*COM MAT*
$\widehat{\mu_\alpha}$	0.0777 (0.0152)	0.0269 (0.0156)	0.0393 (0.0129)	0.0410 (0.00764)	0.0662 (0.0202)	0.00676 (0.00318)
σ_α	0.0457		0.0532	0.0287	0.0803	0.0131
$\widehat{\mu_\beta}$	-0.000262 (0.0000892)	0.000305 (0.000269)		-0.000146 (0.0000383)		
σ_β		0.000734				
$\widehat{\mu_\gamma}$	-0.0211 (0.00948)	-0.00851 (0.0155)		-0.0150 (0.00303)	-0.0172 (0.00303)	
σ_γ						
$\widehat{\mu_\delta}$	0.000320 (0.0000892)	0.000199 (0.000173)		0.000181 (0.0000373)	0.000200 (0.0000232)	
σ_δ						
$\widehat{\phi_1}$	-0.180 (0.00244)		0.620 (0.145)		0.160 (0.0373)	-0.117 (0.0467)
$\widehat{\phi_2}$		0.441 (0.0352)	-0.288 (0.0666)	-0.194 (0.0266)	-0.144 (0.0520)	-0.784 (0.0914)
$\widehat{\phi_3}$			-20.1 (4.39)	4.65 (0.878)		40.0 (17.5)

Table 5: Estimated parameter values and standard errors in parentheses for the best models of hard cost change from July to January for all baskets of repair costs

	RES R&R	*RES LAB*	*RES MAT*	*COM R&R*	*COM LAB*	*COM MAT*
Mean	0.024194	0.051462	0.0066834	0.012046	0.032968	0.0013261
Median	0.017049	0.038829	0.0051755	0.0092433	0.0243	0.0010318
Std. dev.	0.019751	0.038362	0.0052638	0.010916	0.029729	0.0011302
Skewness	1.2319	1.4119	0.59506	0.77689	1.0956	0.7814

Table 6: Statistics of the observed values of σ_{jk}

Distribution	*Param.*	*RES R&R*	*RES LAB*	*RES MAT*	*COM R&R*	*COM LAB*	*COM MAT*
Log-normal	m	-4.2843	-5.0337	-3.3829	-3.9739	-5.5095	-7.3921
	s	1.7245	1.382	1.3025	1.3158	1.2884	2.1754
Pareto	k	-0.31957	-0.27659	-0.31352	-0.19903	-0.59605	-0.59512
	s	0.032175	0.015582	0.067192	0.039683	0.011181	0.0022889
Weibull	a	0.024951	0.012003	0.055151	0.033497	0.0069897	0.0012921
	b	1.1077	0.99108	1.2977	1.0429	1.1517	0.92693

Table 7: Estimated parameter values for parametric distributions of σ_{jk}

For each of the seventeen groups in our available data, the multilevel regression calculates σ_{jk}. Table 6 presents statistics of the observed σ_{jk}, specifically the sample mean, median, standard deviation, and skewness. Figure 19 shows the observed distributions of σ_{jk}, as well as three parametric distributions—log-normal, Pareto, and Weibull—with parameter values estimated from the observed σ_{jk}-distributions. The parametric probability density functions are:

$$\begin{aligned}
f_{\text{log-normal}}\left(\sigma_{jk};\ m,s\right) &= \frac{1}{\sigma_{jk}s\sqrt{2\pi}}\exp\left[\frac{-\left(\log_e\sigma_{jk}-m\right)^2}{2s^2}\right];\\
f_{\text{Pareto}}\left(\sigma_{jk};\ k,s\right) &= \begin{cases} k\frac{s^k}{\left(\sigma_{jk}\right)^{k+1}} & \text{for } \sigma_{jk}>s,\\ 0 & \text{for } \sigma_{jk}<s; \text{and}\end{cases}\\
f_{\text{Weibull}}\left(\sigma_{jk};\ a,b\right) &= \begin{cases} \frac{b}{a}\left(\frac{\sigma_{jk}}{a}\right)^{b-1}\exp\left(\frac{-\sigma_{jk}}{a}\right)^b & \text{for } \sigma_{jk}\geq 0,\\ 0 & \text{for } \sigma_{jk}<0.\end{cases}
\end{aligned}$$

The left-hand side of these equations is read, for example: "the Weibull probability density function, f, is a function of σ_{jk} with parameters a and b." Table 7 lists the estimated parameter values for the log-normal, Pareto, and Weibull distributions applied to the observed σ_{jk}-distributions. We know of no theoretical reason to prefer one parametric distribution of σ_{jk} over another. We find estimated parameter values of these distributions in order to simulate the distributions of future cost changes.

3.4 Simulation of Future Cost Changes

In the previous sections, we found the best models, among those proposed, for the cost change of each basket of repair costs (Equations 5 through 10). We now use these models to simulate cost changes given at least one proximate storm at any United States city along the Atlantic or Gulf coasts in any future Atlantic hurricane season. (Note that the models do not predict the cost change in the absence of a proximate storm.) Needless to say, these simulations are based on the experience of only seven past hurricane seasons. Expectations of future cost changes based on these models are valid only if we assume: that the underlying economic processes resulting in repair costs are stationary with respect to year and region of the Southeastern United States; and that the observed variability of data from these seven seasons is representative of the variability for all future hurricane seasons.

We simulate the cost change of each basket over a range of gradient wind speeds and number of proximate storms. For each simulation, we first sample from the distributions of the stochastic parameters. These parameters, estimated values, and assumed distributions are defined in Equations 5 through 10 and Tables 5 and 7. We assume a Pareto distribution for σ_{jk} in these simulations. When cost changes in the first half of the year appear in the model, we use the median values from our available data: $\Delta_1 = 0.004235$ (RES R&R); $\Delta_2 = 0.009674$ (RES LAB); $\Delta_1 = 0.001310$, $\Delta_2 = 0.01218$, and $\Delta_1\Delta_2 = 0.00008100$ (RES MAT); $\Delta_2 = -0.0004080$ and $\Delta_1\Delta_2 = 0.000007000$ (COM R&R); $\Delta_1 = 0.009926$ and $\Delta_2 = 0.01527$ (COM LAB); and $\Delta_1 = -0.0008180$, $\Delta_2 = -0.0004080$, and $\Delta_1\Delta_2 = 0.000007000$ (COM MAT). We consider gradient wind speeds in the range 10–180 km/hr (in increments of 5 km/hr) and for one to five proximate storms (in increments of 0.1 storms). For each combination of wind speed and number of proximate storms, we simulate the cost change 1000 times. Figures 20 and 21 plot the mean and standard deviation of the

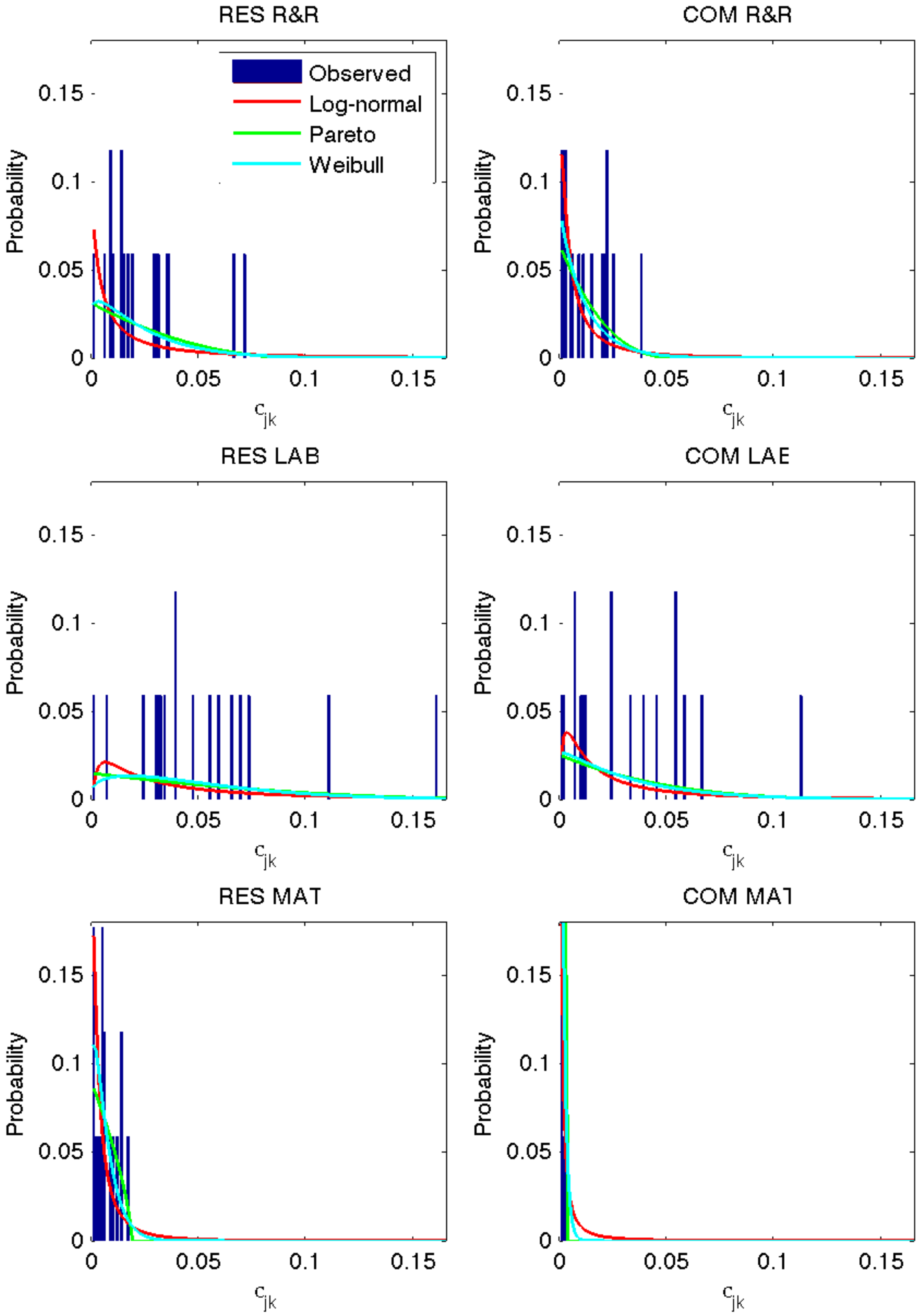

Figure 19: Observed σ_{jk}-distributions and log-normal, Pareto, and Weibull distributions fit to the observations

1000 simulations evaluated at each combination of wind speed and number of proximate storms for the residential and commercial baskets of repair costs, respectively.

The means and standard deviations in Figures 20 and 21 vary with wind speed and the number of storms as a result of the assumed functional forms of the best models. For higher wind speeds and more proximate storms, the cost changes of the remove & replace and labor component baskets are largest. The cost changes of the material component baskets are constant with wind speed and the number of proximate storms. The uncertainties of the simulated cost changes are essentially constant for the remove & replace and material components baskets (and the commercial labor component basket), but the uncertainties of the simulated cost changes of the residential labor component basket increase with increasing wind speed. The cost change of the residential remove & replace basket varies from roughly 0.01 to 0.25, depending on wind speed and the number of proximate storms. The uncertainty about any specific cost change prediction (two standard errors of prediction) is roughly 0.1. The cost change of the commercial remove & replace basket varies from roughly 0.005 to 0.15, with an uncertainty of roughly 0.08.

Finding 19 *Assuming median values for the cost changes in the first half of the year, the best model simulates cost changes of the residential remove & replace basket between 0.01 and 0.25, depending on the given wind speed and number of proximate storms, with two standard errors of prediction equal to 0.1. The best model simulates cost changes of the commercial remove & replace basket between 0.005 and 015, with two standard errors of prediction equal to 0.08.*

Finding 20 *The uncertainty in simulated values of the cost change of the residential labor component basket increases with increasing wind speed, but the uncertainty in simulated values of the five other baskets are essentially constant with wind speed and the number of proximate storms.*

We can also identify whether the uncertainties given wind speed and the number of proximate storms result from differences between groups or from unexplained variation. For example, consider the model for the residential remove & replace basket of repair costs (repeated here):

$$\Delta_i^{\text{RES R\&R}} = \alpha_{jk} + \beta w_i + \gamma n_{ci} + \delta w_i n_{ci} + \phi_1 \Delta_{1i} + \epsilon_{jk}.$$

The variance is:

$$\begin{aligned} \operatorname{Var}\left[\Delta_i^{\text{RES R\&R}}\right] &= \operatorname{Var}\left[\alpha_{jk} + \epsilon_{jk}\right] \\ &= \operatorname{Var}\left[\alpha_{jk}\right] + \operatorname{Var}\left[\epsilon_{jk}\right]. \end{aligned}$$

In the simulations shown in Figure 20, the variances for simulations of the residential remove & replace basket cost changes range from 0.00210 to 0.00282, depending on wind speed and the number of storms. Returning to the original data, the variance of the observed cost changes of the residential remove & replace basket is 0.00536. Thus, our best model accounts for 47.4 to 60.8 percent of the variance in the data. The variance of α_{jk}, σ^2_α, ranges from 0.00178 to 0.00241, which is consistent with $\sigma^2_\alpha = 0.00209$ in Table 5. The variance of ϵ_{jk}, σ^2_{jk}, ranges from 0.000310 to 0.000413, which is consistent with $\sigma^2_{jk} = 0.000390$ in Table 6. Of the variance in a prediction of the cost change of the residential remove & replace basket, 82 to 87 percent is from the variance of the intercept.

We repeat the above analysis for the remaining five baskets. From the simulations shown in Figures 20 and 21, the variances for predictions of cost change are: 0.000137–0.0189 (RES LAB); 0.00387–0.00503 (RES MAT); 0.00134–0.00186 (COM R&R); 0.00560–0.00753 (COM LAB); 0.000146–0.000207 (COM MAT). The variances of the original data and the percentages of these variances accounted for in the model are: 0.0201, 6.11–99.3 percent (RES LAB); 0.00259 and the variance of the model exceeds the variance of the data by 49.4–94.2 percent (RES MAT); 0.00227, 18.0–40.9 percent (COM R&R); 0.0176, 57.1–68.1 percent (COM LAB); 0.000140 and the variance of the model exceeds the variance of the data by 4.29–47.9 percent (COM MAT). Generally speaking, our models for the remove & replace and labor component baskets account for about half of the variance of the observed data, and our models for the material baskets have a larger variance than that of the observed data.

Finding 21 *When we use the best models to simulate future cost changes of the remove & replace and labor component baskets, the variances of the simulated data are roughly half of the variances of the observed data. There is more observed variability in the remove & replace and labor component data than we can simulate with the best models. For the material component baskets, the variances of the simulated data exceed those of the observed data.*

The following numbers are the percentages of the variance in a prediction of cost change from the variance of the intercept (or slope in the RES LAB model): 33.0–99.5 (RES LAB); 59.1–68.3 (RES MAT); 45.6–56.7 (COM R&R); 99.6–99.7 (COM LAB); 99.3–99.5 (COM MAT). Most of the variance in a predicted value comes from uncertainty between groups by U.S. region and year, not from uncertainty within the groups. This group-to-group uncertainty remains after our models account for wind speed, the number of hurricanes, and cost changes in the first half of the year.

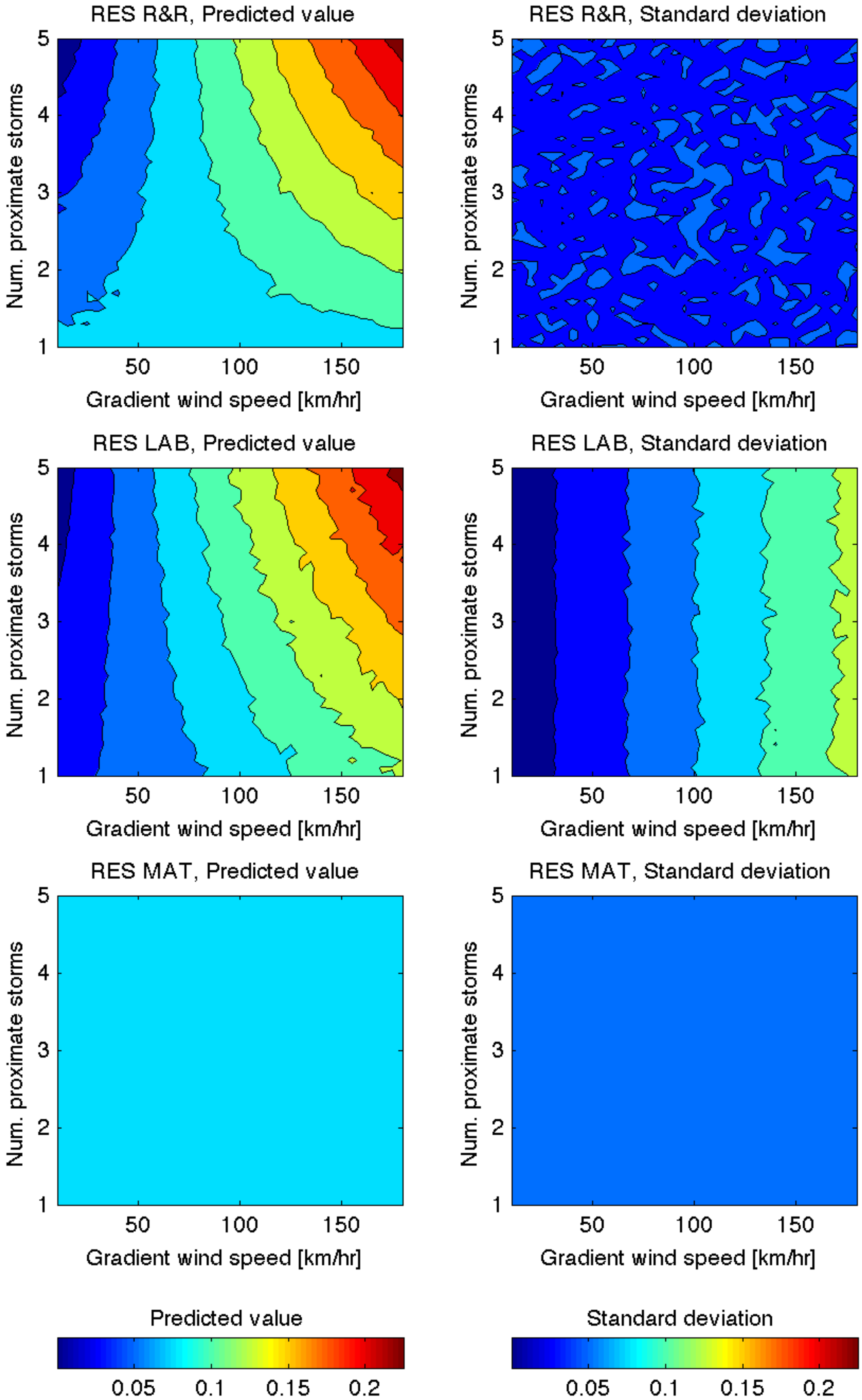

Figure 20: Simulations of cost changes from July to January for the residential baskets of repair costs. Plots on the left show the expected value of the simulation as a function of gradient wind speed and the number of proximate storms from 1000 simulations at each combination of wind speed and number of storms. Plots on the right show the standard deviations of these simulations.

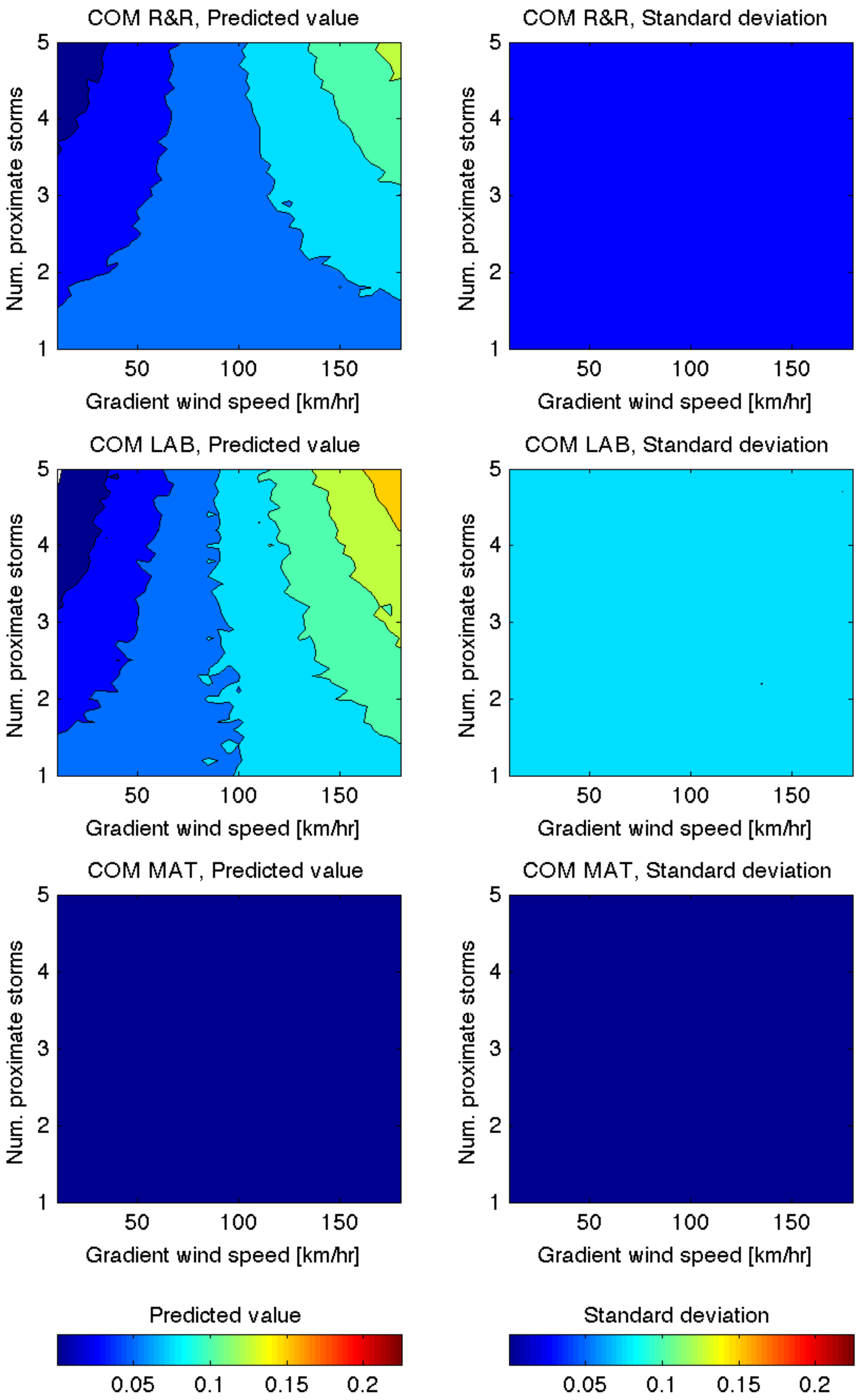

Figure 21: Simulations of cost changes from July to January for the commercial baskets of repair costs. Plots on the left show the expected value of the simulation as a function of gradient wind speed and the number of proximate storms from 1000 simulations at each combination of wind speed and number of storms. Plots on the right show the standard deviations of these simulations.

Finding 22 *The primary source of uncertainty in a prediction from our best models is the stochastic parameters, not the residual variation. If we know the parameter values for a given group with certainty, we can make a relatively precise prediction of the cost change. However, since the parameter values are so uncertain for a particular region and year, a prediction of cost change for that group is relatively imprecise.*

4 Discussion

Although the data used in this study are the best available to us, they are not ideal. In most years, the Xactimate data are reported at the beginning of each quarter; data since 2008 are reported at the beginning of each month. The Atlantic hurricane season is officially from 1 June to 30 November, but most hurricanes happen in August, September, and October. One (or more) large tropical storm could affect a city at any time between reports of cost data in July and the following January. Thus our data reflect economic processes independent of *and* resulting from a hurricane. For example, there may be a shortage of materials in the months before a hurricane affects a city, but we cannot resolve the effects of the shortage from those of the hurricane. These issues presumably explain some of the variance in our data and residual variance of our models.

The way we calculate wind speed at a city certainly contributes to the variance in our models. When we initially studied the material price and labor wage data, we used the wind speeds along the path of the storms. We improved this approach by calculating the wind speed at the city from observations of the storm system via the Holland Wind Profile (Holland, 1980). Better estimates of wind speeds at a city would certainly reduce some of the uncertainty in these models.

Modeling demand surge is a multivariate problem with an imprecise response variable and unknown explanatory variables. We select a part of the demand surge phenomenon, precisely define a response variable, and compare models with different sets of explanatory variables. The available data suggest that wind speed, the number of tropical storms affecting a city, the year, and the state in which a city is located may contribute to how costs change after large tropical storms. These variables are presumably proxies for more fundamental contributors to cost changes. Wind speed and the number of storms are likely proxies for the amount of damage to properties. The year may be a proxy for underlying economic processes affecting demand for construction materials and labor before the hurricane season. The state may represent local- or state-level governmental regulations regarding the movement of materials and labor after a natural disaster. The use of different, additional, or more refined explanatory variables may help to reduce the variance of our best models.

Despite their limitations, our best models provide insight into one aspect that is believed to be the largest contributor to demand surge. The use of multilevel regression models allows us to test for the significance of explanatory variables in a proposed model and identify how much variation in the changes of costs is due to within-group versus between-group differences. By proposing several models for the cost change of each basket of repair costs, we can show that one model accounts for the observed data better than another. For the material component baskets, our best models do not include variables for the wind speed and number of storms; the cost changes in the first half of the year are sufficient to account for cost changes in the second half of the year. The best models for the cost changes of the remove & replace and labor component baskets do include variables for both the wind speed and number of storms, as well as cost changes in the first half of the year.

The models for the remove & replace and labor component baskets account for roughly half of the variability in the observed data. Wind speed, the number of proximate storms, and cost changes in the first half of the year do not account for all the variability in the observed data. This suggests that either the explanatory variables we tried must be refined (for example, by using wind speeds observed at cities not inferred from the storm systems) or we must use explanatory variables in addition to, or instead of, the ones we tried. Given our best models, most of the uncertainty in a simulation of cost changes is due to differences between groups defined by U.S. geographic region and year, not differences within a given region and year. The between-group uncertainty might be reduced by introducing a group-level variable, such as the number of tropical storms making landfall in a given region and year or the demand for new construction in the given region and year.

For the material component baskets, the models have more uncertainty than the observed data. In other words, we could predict a future cost change of the material component basket by sampling from the empirical distribution defined by observed cost changes, and the uncertainty in this prediction is less than the uncertainty in a prediction from our best models. We do not recommend the empirical model because the regression analyses suggest that the cost changes in the first half of the year are significant in predicting the cost change in the second half of the year. Moreover, our best regression models allow for variability in future observations suggested, but not seen, in past observations.

There may be situations where a prediction of cost changes during an Atlantic hurricane season is desired, but there are no available data on the cost changes in the first half of the year. For example, reinsurance contracts are typically renewed in the last two months of a calendar year, so basket cost changes in the next two quarters cannot be known. Since our best models all include these explanatory variables, the models cannot be used here in the way they are intended. There are at least two ways to make a prediction nonetheless. One possibility is to use median values for these cost changes (as presented in Section 3.4) or to use a prediction of the cost changes. These values would be used in the models as if they were known, however, the uncertainties in the resulting predictions would be lower bounds. The actual uncertainty in such a prediction should be larger than the uncertainty if known values of cost changes were used. A second possibility is to find a model that does not use cost changes in the first half of the year as explanatory variables. For the residential, remove & replace and labor component baskets, we proposed models of this type (see Tables 4 and B.8) but found that different models have lower AIC values. Models for the commercial, remove & replace and labor component baskets could be found in a similar manner. For the material component baskets, predictions of cost changes during an Atlantic hurricane season could be made directly from the observed, empirical distributions of the cost changes from the past nine seasons.

At the beginning of Section 3, we raised several general questions, and we now address them based on the results of this study. On the Atlantic and Gulf Coasts, the most extreme cost changes occur only when there is a large proximate storm (that is, a tropical storm within 200 km of a city with a surface wind speed of at least 50 km/hr inferred from the storm system using the Holland Wind Profile). There can be cost changes of 20 percent or more in the absence of such a storm, but there are only a few such observations. (See Figures 2 and 8.) There are many more observations of large cost changes when there is a large proximate storm than when there is not, and theoretically, we expect large cost changes after such storms. Thus, there is evidence that large proximate storms cause large cost changes. More intense tropical storms and a greater number of storms affecting a city generally cause higher cost changes. (See Figures 20 and 21.) The exception is the cost change of the material component basket, which does not depend on wind speed or the number of proximate storms. This study does not identify any regional or annual factors that might explain differences in cost changes resulting from the region of the Southern United States where a city is located or the year of observation.

5 Conclusions

- Changes in labor costs drive the changes in total repair costs. A better understanding of the labor market, rather than the market for materials, would likely provide a better understanding of why total repair costs increase after natural disasters. For our total repair cost and labor component baskets, the best models to predict cost changes include variables accounting for the wind speed of a tropical storm, the number of tropical storms affecting a city, and the cost changes of the baskets in the first half of the year. All these variables contribute significantly to the prediction.

- For our material component baskets of repair costs, the best models include variables for cost changes in the first half of the year, but the models do not include variables for wind speed or the number of storms affecting the city. Changes in the costs of reconstruction materials do not appear to be affected by Atlantic tropical storms.

- For the remove & replace and labor component baskets, our best models to predict cost changes—using wind speed, the number of storms affecting a city, and cost changes in first half of year—account for roughly half of the observed variability in cost changes during Atlantic hurricane seasons. To explain the remaining half of the observed variability, one might refine these variables or seek additional explanatory variables.

- Given our best models, there is more variability between groups defined by geographic region of the United States (specifically, states on the Gulf of Mexico, Florida, or states on the Atlantic Ocean) and year than variability within a group. Our models could be improved by explaining and modeling these between-group variations.

6 Acknowledgments

This work was sponsored by Willis Group Holdings Limited through the Willis Research Network. We appreciate the support and interest of our colleagues currently or formerly at Willis: Rowan Douglas, Matthew Foote, Julie Serakos, and Kyle Beatty. James Done provided helpful guidance on understanding the Holland Wind Profile and applying it to tropical storms. Also, we appreciate the suggestion of Chris Olsen to use multilevel regression to model the available data.

References

Australian Securities & Investments Commission (2007, January). Making home insurance better. Report 89, Australian Securities & Investments Commission.

Burnham, K. P. and D. R. Anderson (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Second ed.). New York: Springer.

Demuth, J. L., M. DeMaria, and J. A. Knaff (2006, November). Improvement of advanced microwave sounding unit tropical cyclone intensity and size estimation algorithms. *Journal of Applied Meteorology and Climatology 45*(11), 1573–1581. Version 1.10, 16 November 2010.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, New York: Cambridge University Press.

Goldstein, H. (2003). *Multilevel Statistical Models* (Third ed.). Kendall's Library of Statistics. London: Arnold.

Guy Carpenter (2005, September). Model issues to consider in Katrina's wake. Instrat Briefing.

Hallegatte, S. (2008). An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina. *Risk Analysis 28*(3), 779–799.

Holland, G. J. (1980, August). An analytic model of the wind and pressure profiles in hurricanes. *Monthly Weather Review 108*, 1212–1218.

Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications* (Second ed.). New York, New York: Routledge.

International Hurricane Research Center (2009, 22 May). Florida public hurricane loss model. *Technical Report*, Florida International University, Miami, Florida. Submission in compliance with the 2008 Standards of the Florida Commission on Hurricane Loss Projection Methodology. FPHLM version 3.1.

Johnson, R. A. and D. W. Wichern (2002). *Applied Multivariate Statistical Analysis* (Fifth ed.). Upper Saddle River, New Jersey: Prentice Hall.

Kuzak, D. and T. Larsen (2005). Use of catastrophe models in insurance rate making. In P. Grossi and H. Kunreuther (Eds.), *Catastrophe Modeling: A New Approach to Managing Risk*, Chapter 5. New York, NY: Springer.

MathWorks (2010a, 16 August). Boxplot help file. MATLAB R2010b.

MathWorks (2010b, 16 August). Kstest2 help file. MATLAB R2010b.

National Atlas of the United States (2003, 5 March). Map maker. `http://www.nationalatlas.gov`. Accessed 9 August and 27 October 2010.

Okuyama, Y. (2007, June). Economic modeling for disaster impact analysis: Past, present, and future. *Economic Systems Research 19*(2), 115–124.

Olsen, A. H. and K. A. Porter (2010, January). What we know about demand surge. *Technical Report* SESM-10-1, Department of Civil, Environmental, and Architectural Engineering, University of Colorado at Boulder.

Olsen, A. H. and K. A. Porter (2011, May). What we know about demand surge: Brief summary. *Natural Hazards Review 12*(2), 62–71.

Pinheiro, J. C. and D. M. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. New York, New York: Springer-Verlag.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rose, A. (2004). Economic principles, issues, and research priorities in hazard loss estimation. In Y. Okuyama and S. E. Chang (Eds.), *Modeling Spatial and Economic Impacts of Disasters*, Chapter 2, pp. 13–36. New York: Springer-Verlag.

Xactware (2008). Retail labor rates and supporting events. Pricing white paper on website.

Xactware (2009, 10 April). Pricing research methodology. White paper.

Xactware (2011a). About Xactware. `http://www.xactware.com/company/about`. Accessed 16 February 2011.

Xactware (2011b). February 2011 price list update now available. `http://www.xactware.com/news/release20110201`. Accessed 16 February 2011.

Xactware (2011c). Replacement cost estimating with Xactimate 27. http://www.xactware.com/solutions/claims-estimating/. Accessed 16 February 2011.

Xactware (2011d). Xactimate pricing data services. http://www.xactware.com/solutions/claims-estimating/#pricing-data-services. Accessed 16 February 2011.

Xactware (2011e). Xactware's pricing data. http://www.xactware.com/resources/pricing-data-services/. Accessed 16 February 2011.

Xactware (2011f). Xactware's pricing research and methodology. http://www.xactware.com/resources/pricing-data-services/#pricing-methodology. Accessed 16 February 2011.

A Complete Descriptions of Baskets from Xactimate Price Lists

(Tables begin on the next page.)

Short description	*Abbreviation*	*Amount*	*Remove description*	*Replace description*
Composition shingle roofing	RFG220	300 sf	"Includes: labor cost to remove shingles and discard in a job site waste receptacle."	"Includes: 3 tab composition shingles, 15 pound roofing felt, roofing nails, and installation labor. Quality: 3 tab with a 20 year warranty, and a class A fire rating."
Clay tile roofing	RFGTIL	100 sf	"Includes: labor cost to remove clay tile roofing and discard in a job-site waste receptacle."	"Includes: tile roofing, furring strips, 30 pound roofing felt, roofing nails, and installation labor. Quality: clay 'S' or flat tile."
Reglazing	WDR1/4	40 sf	n/a	"Includes: clear glass pane, rubber gasket or window caulk, and installation labor. Quality: up to 40 sf of 1/4 in clear glass."
Installed drywall	DRY5/8	100 sf	"Includes: labor cost to remove 5/8 in drywall and discard in a job-site waste receptacle."	"Includes: drywall, metal corner bead, joint compound ('mud'), drywall nails, perfotape, grabber screws, texture machine, and installation labor. Quality: 5/8 in type 'X' drywall with light 'orange peel' texture, ready for paint."
Carpet	FCCAV	100 sf	"Includes: labor cost to remove carpet, tackless strip, and seaming tape and discard in a job-site waste receptacle."	"Includes: carpet, tackless strip, seaming tape, seaming iron, power carpet stretcher, and installation labor. Excludes: metal transition strip. ... Quality: polyester or polypropylene carpet with average pile density."
Sealed and painted wood siding	PNTSDG	200 sf	n/a	"Includes: latex primer, latex paint, acrylic caulk, painter's putty, sandpaper, and labor."

Table A.1: Complete descriptions of material and labor items, and their amounts, in the residential baskets of repair costs.

Short description	*Abbreviation*	*Amount*	*Remove description*	*Replace description*
3-ply roofing	RFGBU3	100 sf	"Includes: labor cost to remove built-up roofing and discard in a job-site waste receptacle."	"Includes: 1 [one] felt base sheet and 2 felt ply sheets, gravel stop, roofing nails, hot asphalt (tar), and installation labor."
Gravel ballast roofing	RFGBUG	100 sf	"Includes: labor cost to remove gravel ballast and discard in a job-site waste receptacle."	"Includes: pea gravel, the use of a conveyor belt, and installation labor."
PVC membrane roofing	RFGSPLY	100 sf	"Includes: labor cost to remove a PVC membrane and discard in a job site waste receptacle."	"Includes: PVC membrane, seam adhesive, fasteners, and installation labor."
Metal roofing	RFGMTL	100 sf	"Includes: labor cost to remove metal roofing and discard in a job-site waste receptacle."	"Includes: metal roofing sheets, roofing screws with neoprene washer, and installation labor. Quality: baked-on color finish, ribbed overlapping, 29 gague roofing."
Glass curtain wall	GLSCW	100 sf	"Includes: labor cost to remove a single glazed curtain wall and discard in a job-site waste receptacle."	"Includes: glass, hardware, and installation labor."
Installed drywall	DRY5/8	100 sf	(See Table A.1.)	(See Table A.1.)
Carpet	FCCAV	100 sf	(See Table A.1.)	(See Table A.1.)
Sealed and painted wood siding	PNTSDG	200 sf	n/a	(See Table A.1.)

Table A.2: Complete descriptions of material and labor items, and their amounts, in the commercial baskets of repair costs.

B Additional Figures and Tables

	p-value						
Basket	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*
RES R&R	0.3232	0.2734	0.0010	0.0002	0.1740	0.0438	0.0000
COM R&R	0.2000	0.7448	0.0002	0.0000	0.4794	0.0438	0.0001
RES LAB	0.4894	0.5931	0.0005	0.0002	0.2625	0.1622	0.0001
COM LAB	0.1167	0.5931	0.0000	0.0001	0.1895	0.5697	0.0001
RES MAT	0.1120	0.0800	0.0038	0.5170	0.3417	0.0025	0.0006
COM MAT	0.0236	0.0020	0.0173	0.9903	0.2947	0.0025	0.0006

Table B.1: P-values from hypothesis tests that observed market cost changes in a given year do not differ between cities affected and unaffected by large proximate storms

	p-value								
Basket	*TX*	*LA*	*MS*	*AL*	*FL*	*GA*	*SC*	*NC*	*VA*
RES R&R	0.0006	0.0002	0.0002	0.0003	0.0000	0.4806	0.0392	0.0860	0.9350
COM R&R	0.0000	0.0007	0.0028	0.0003	0.0000	0.4806	0.0707	0.0058	0.5344
RES LAB	0.0006	0.0274	0.0098	0.0002	0.0000	0.1994	0.3950	0.0642	0.9350
COM LAB	0.0000	0.0532	0.0799	0.0012	0.0000	0.1052	0.0562	0.9012	0.5344
RES MAT	0.0000	0.0000	0.6693	0.1766	0.0000	0.1994	0.0236	0.0058	0.5344
COM MAT	0.0020	0.0371	0.6693	0.7096	0.0000	0.1626	0.0158	0.0001	0.1984

Table B.2: P-values from hypothesis tests that observed market cost changes in a given U.S. state do not differ between cities affected and unaffected by large proximate storms

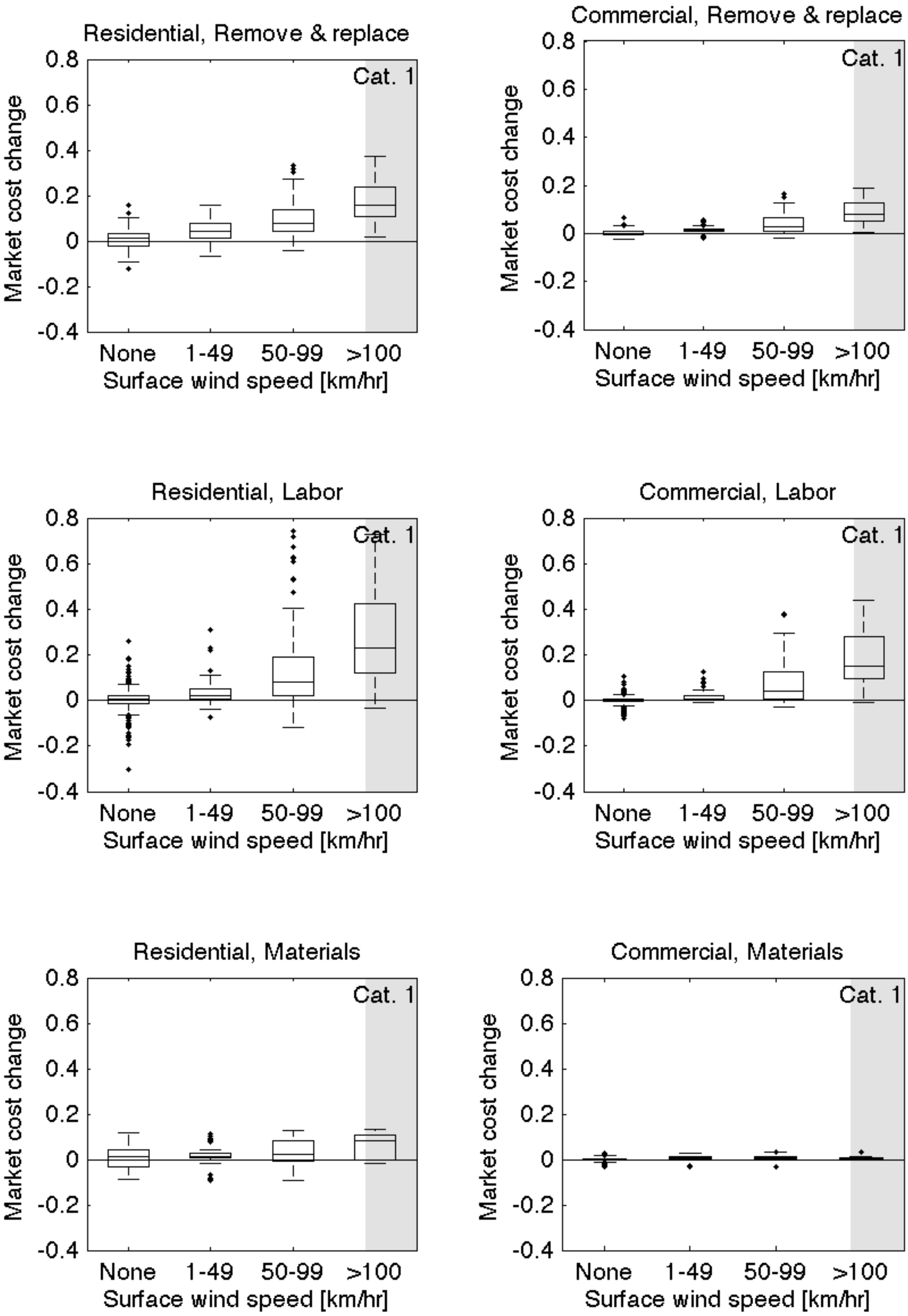

Figure B.1: Observed market cost change from July to January categorized by wind speed

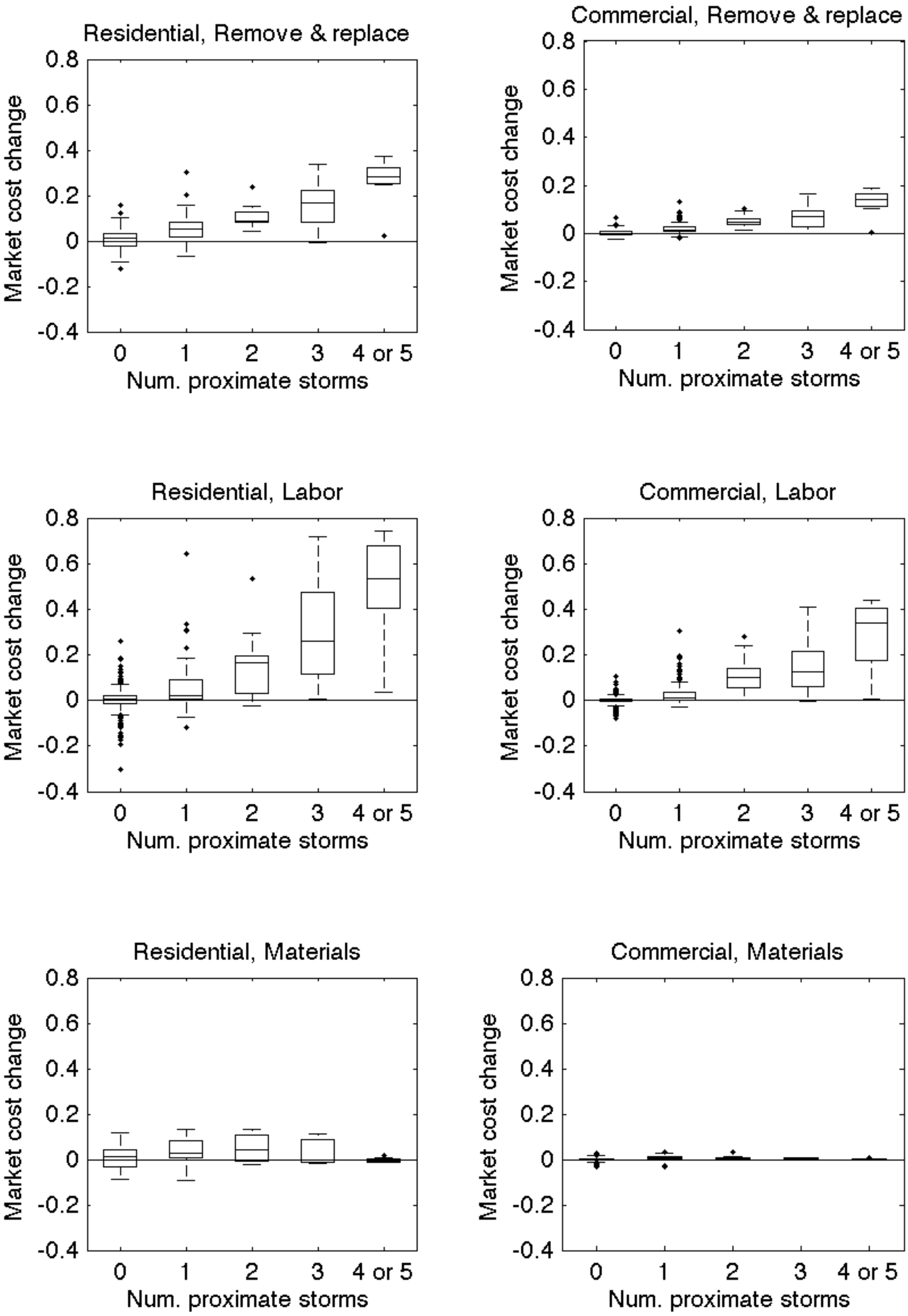

Figure B.2: Observed market cost change from July to January categorized by the number of proximate storms

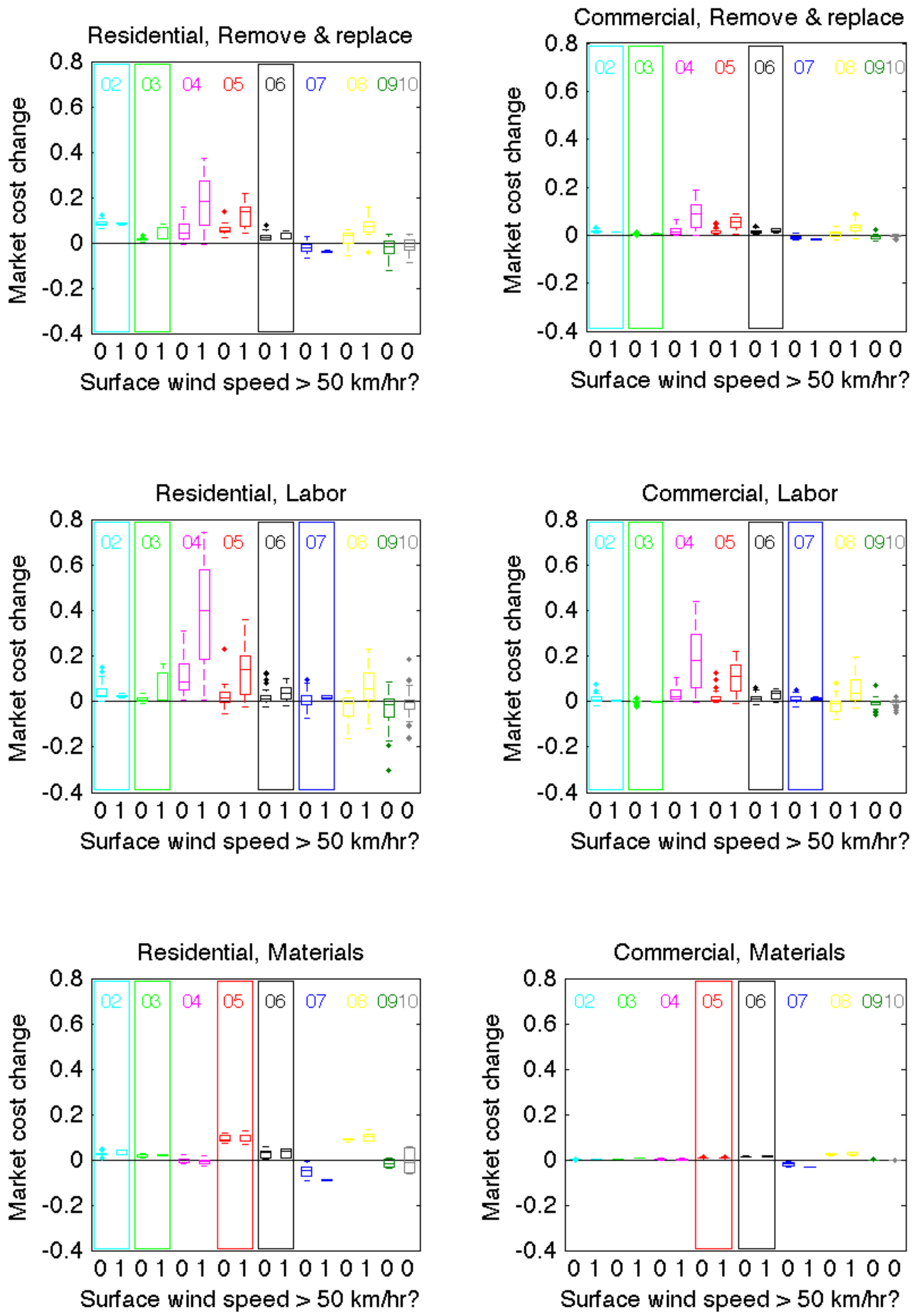

Figure B.3: Observed market cost change from July to January categorized by year. Rectangles around box-and-whiskers indicate that we could not reject the null hypothesis that, in a given year, cost changes in cities affected by a large proximate storm are drawn from the same population as cost changes in cities unaffected by such a storm.

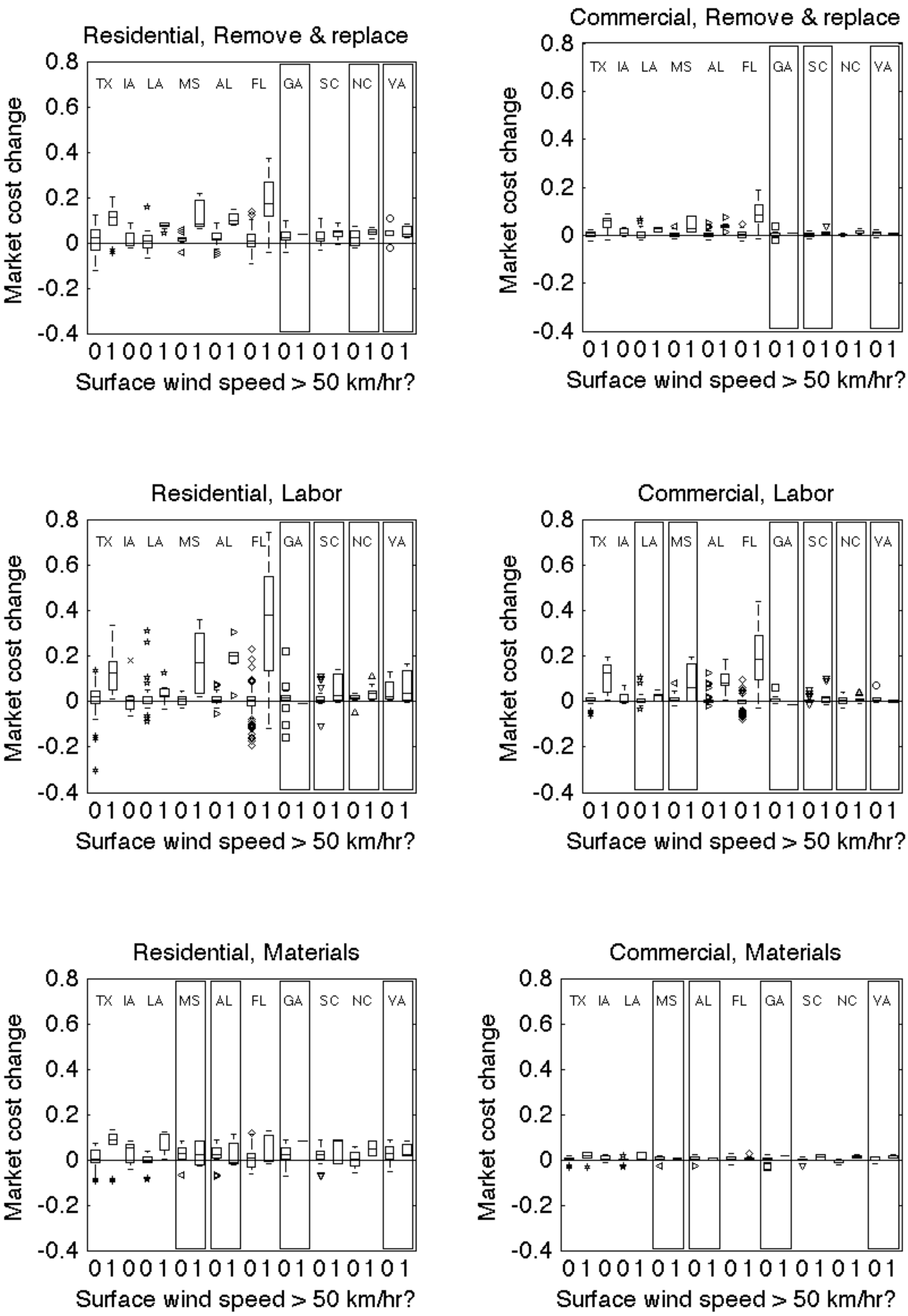

Figure B.4: Observed market cost change from July to January categorized by U.S. state. Rectangles around box-and-whiskers indicate that we could not reject the null hypothesis that, in a given U.S. state, cost changes in cities affected by a large proximate storm are drawn from the same population as cost changes in cities unaffected by such a storm.

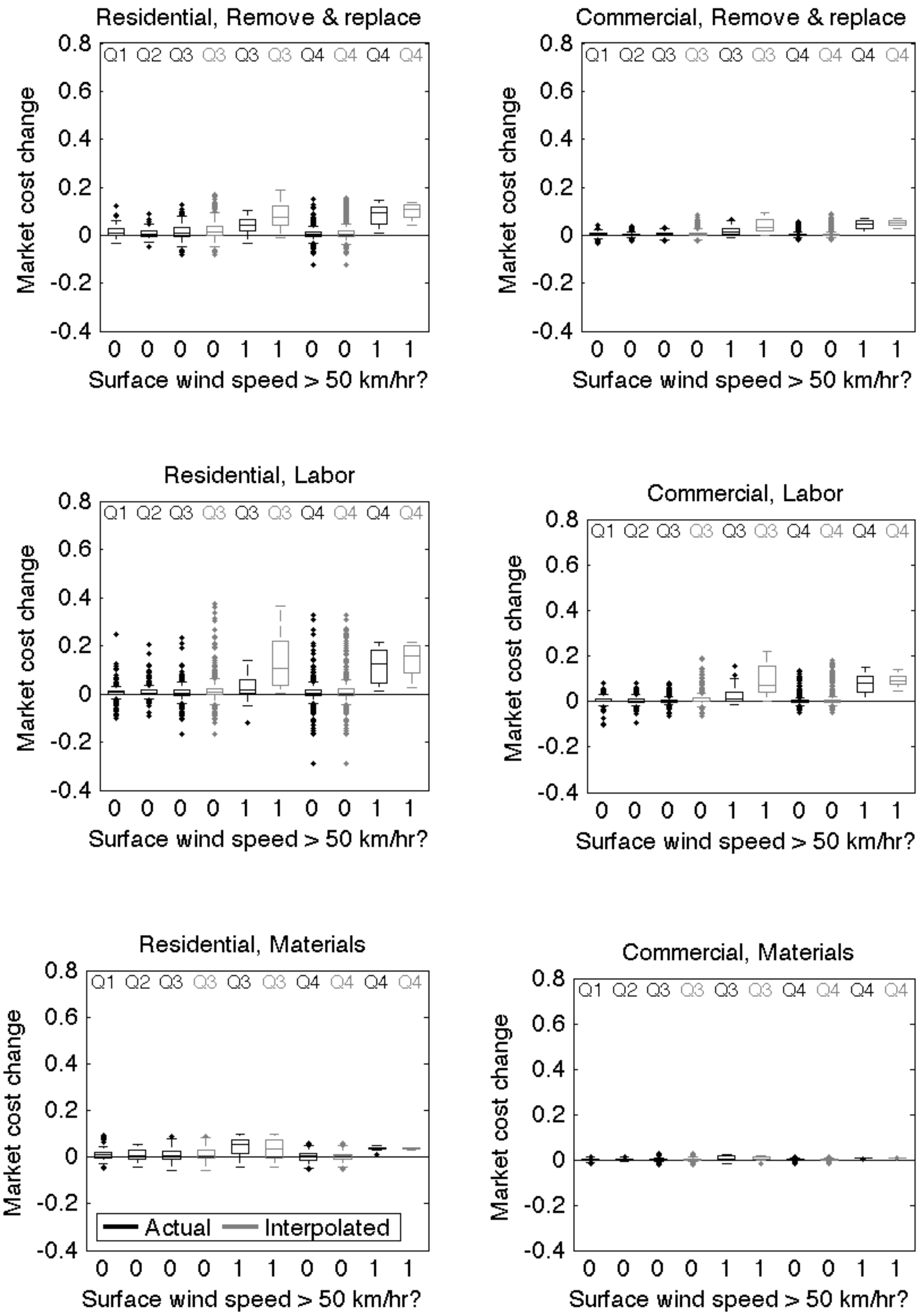

Figure B.5: Observed market cost change by quarter. When there are missing data in the Xactimate databases, we linearly interpolate the missing datum from the preceding and succeeding quarters. The data as extracted from the Xactimate databases form the "actual" data sets, and these data sets augmented with the linearly interpolated points form the "interpolated" data sets.

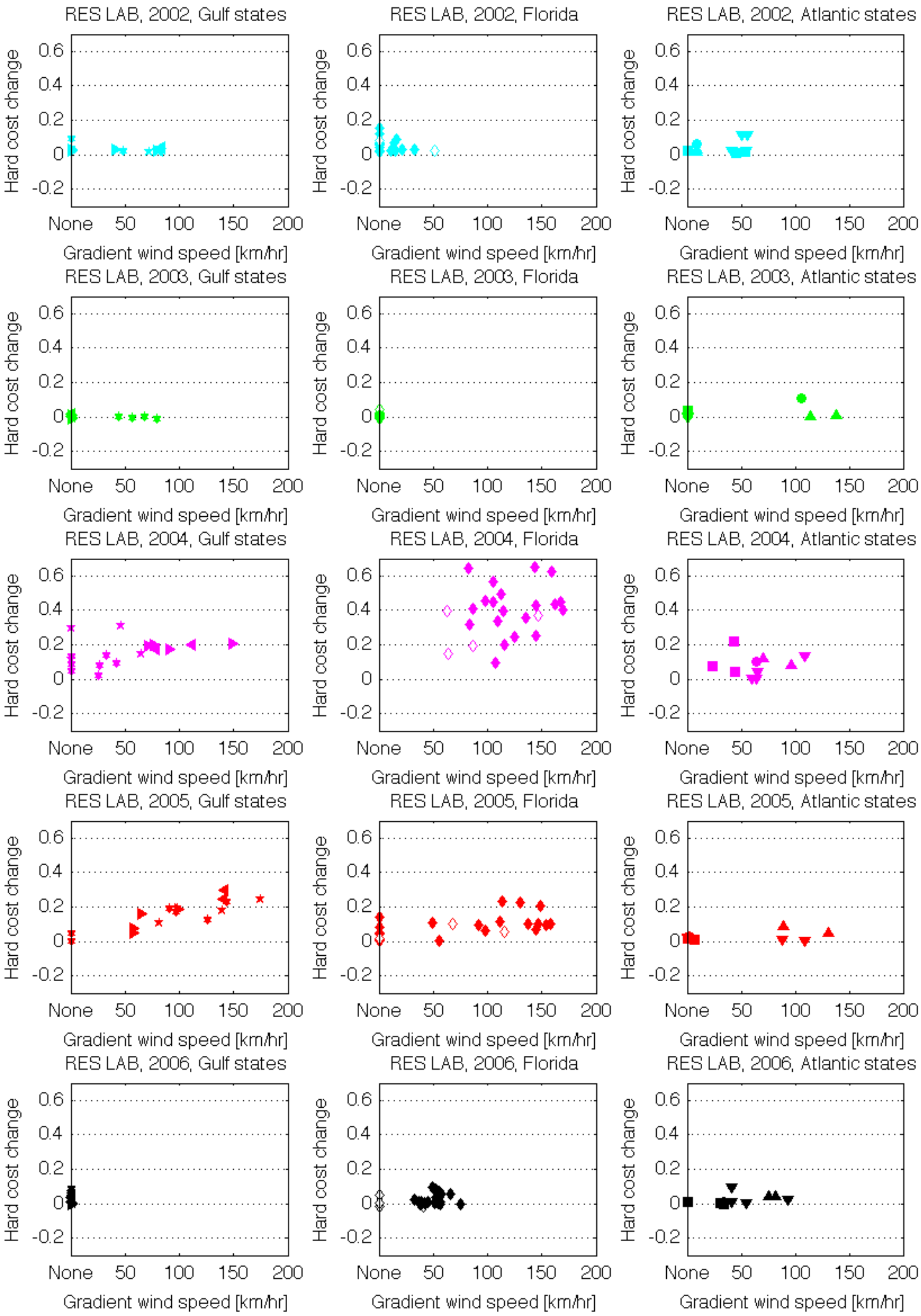

Figure B.6: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the residential labor component basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

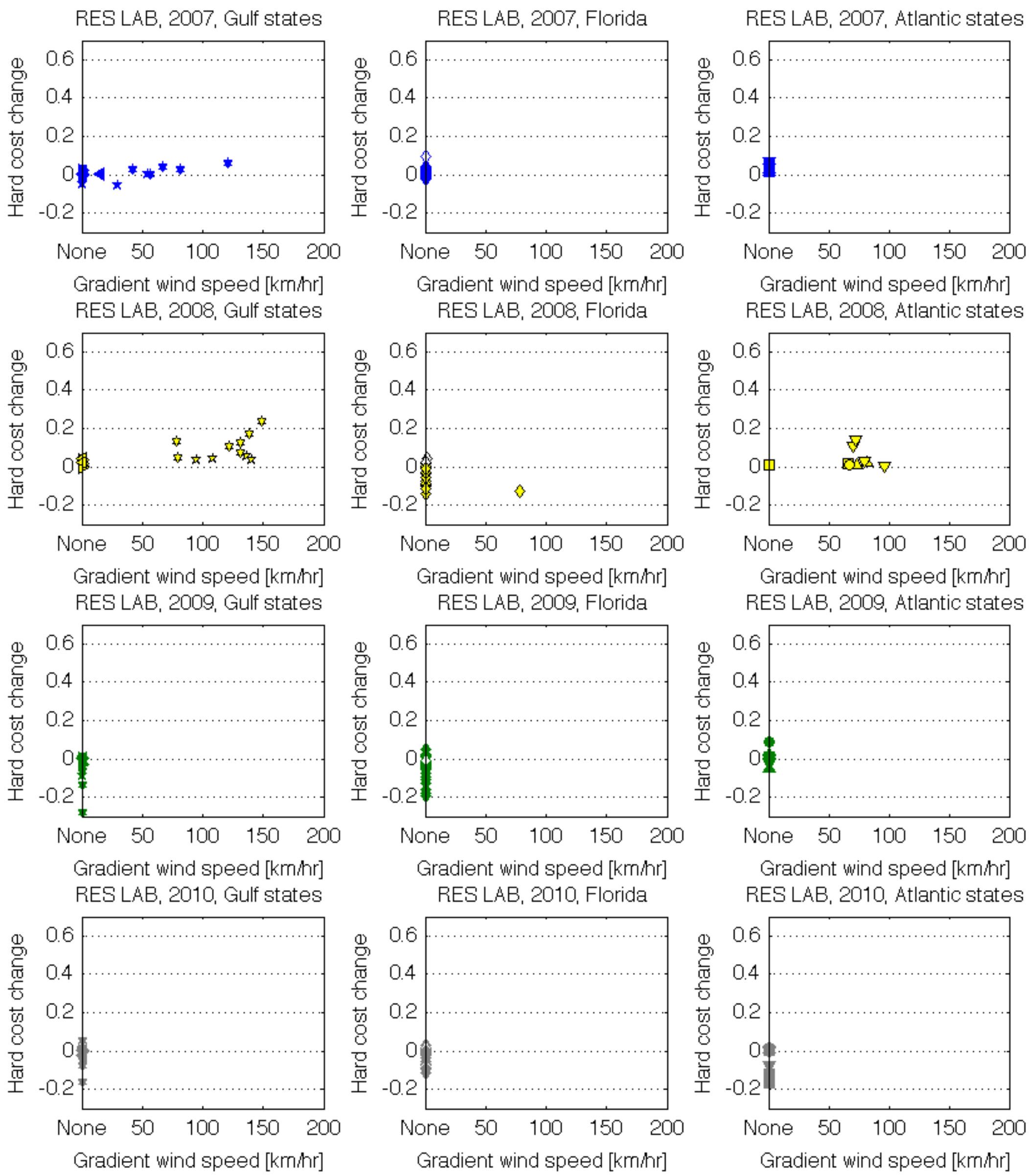

Figure B.7: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the residential labor component basket.

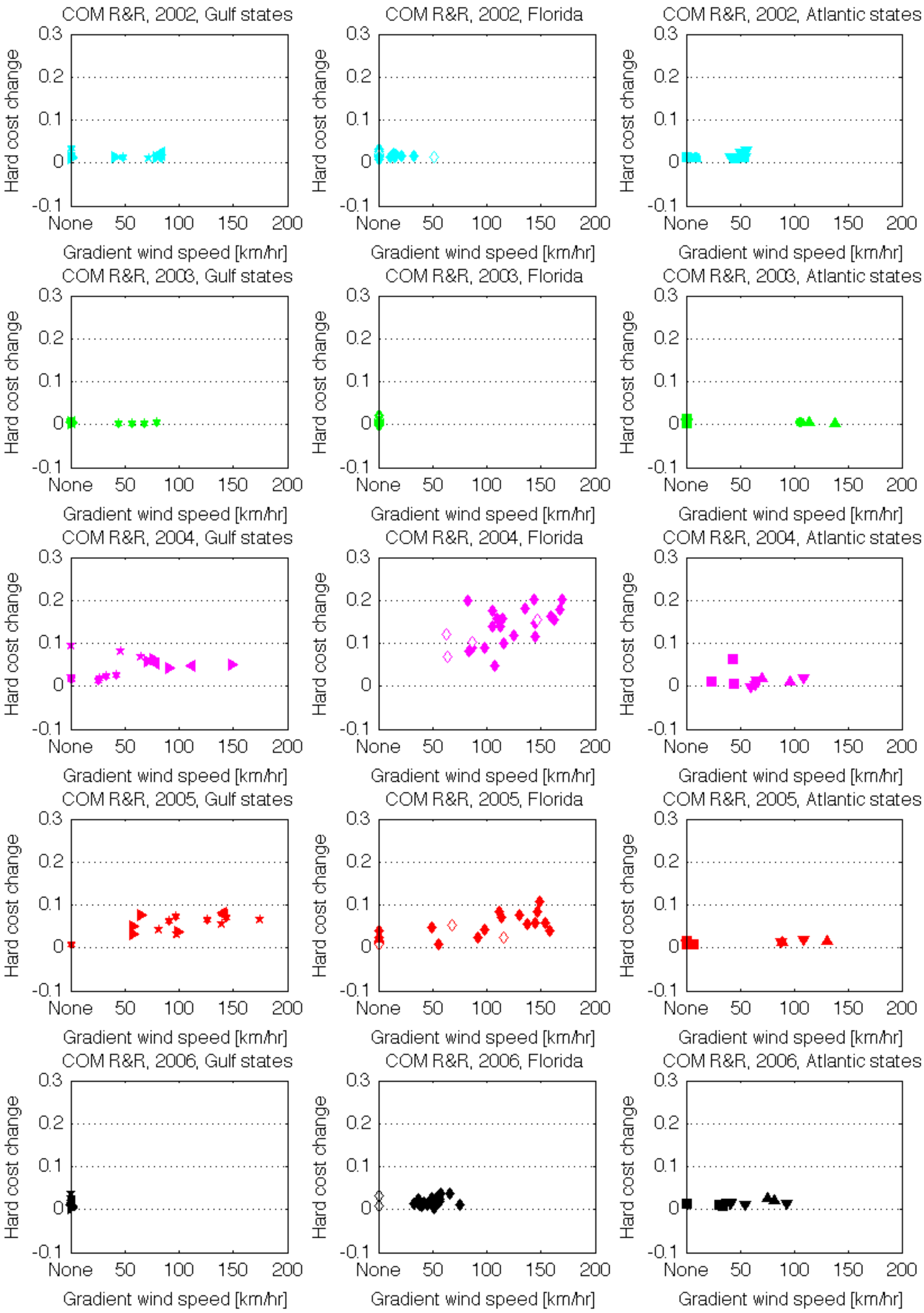

Figure B.8: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the commercial remove & replace basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

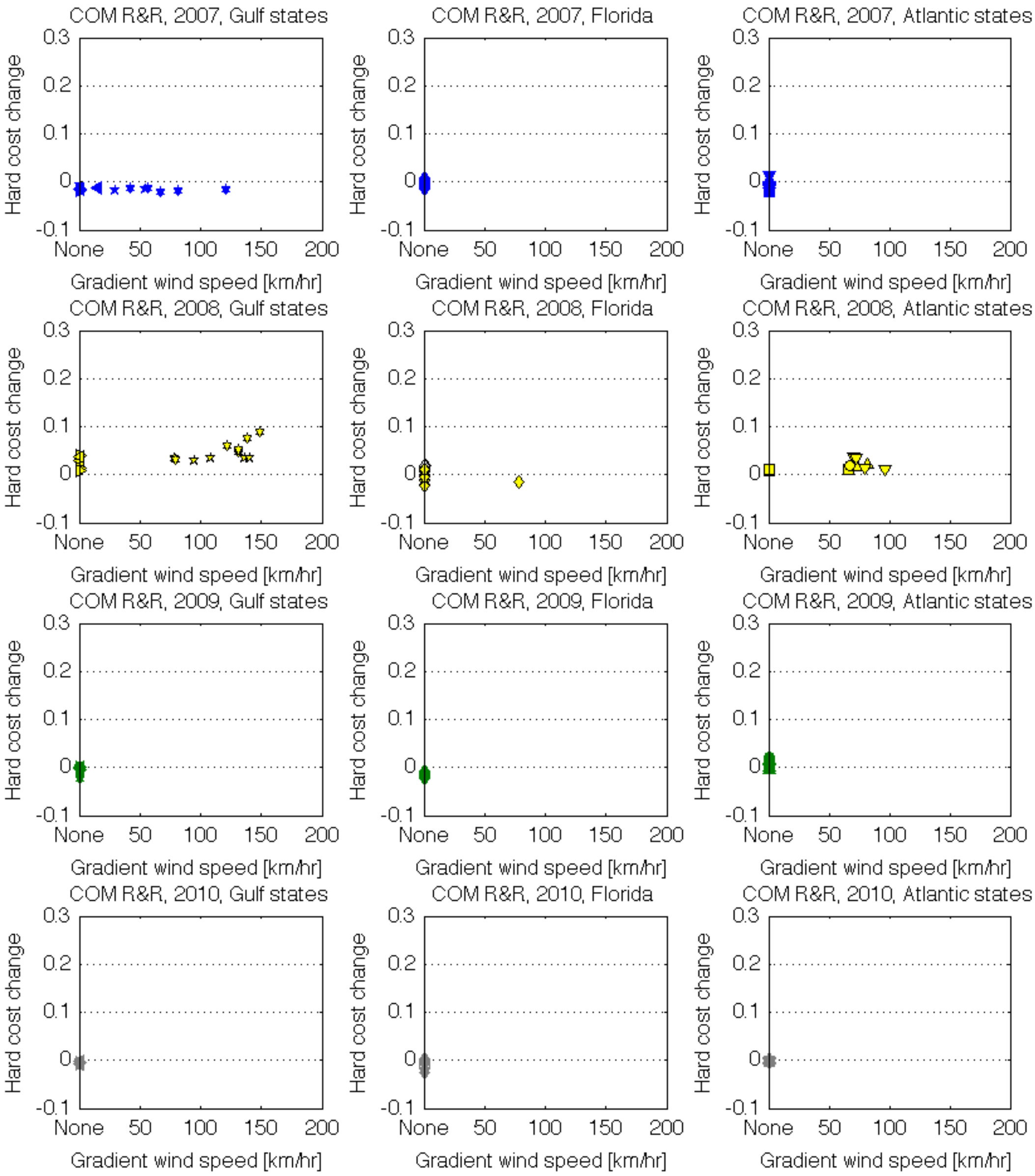

Figure B.9: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the commercial remove & replace basket.

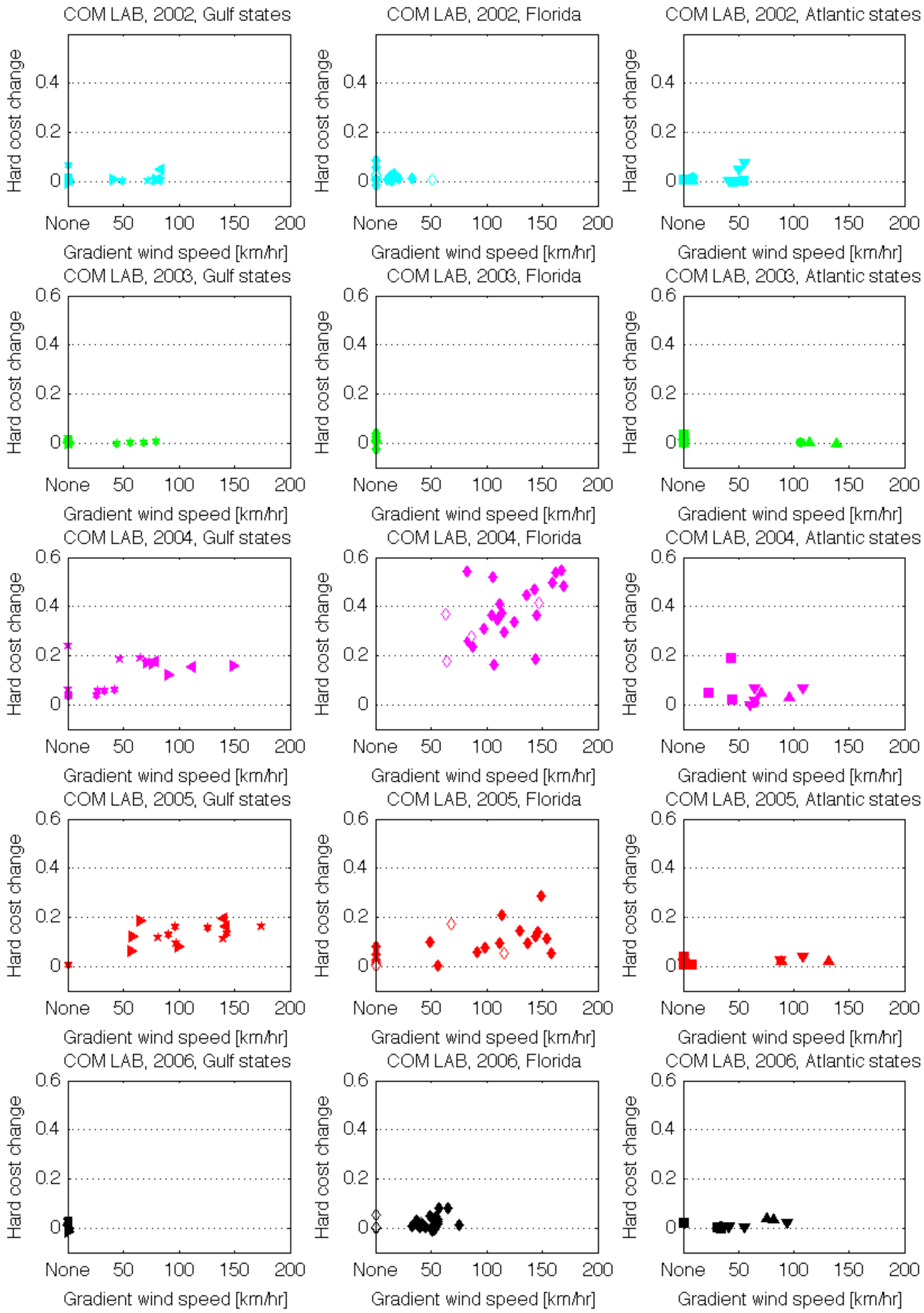

Figure B.10: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the commercial labor component basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

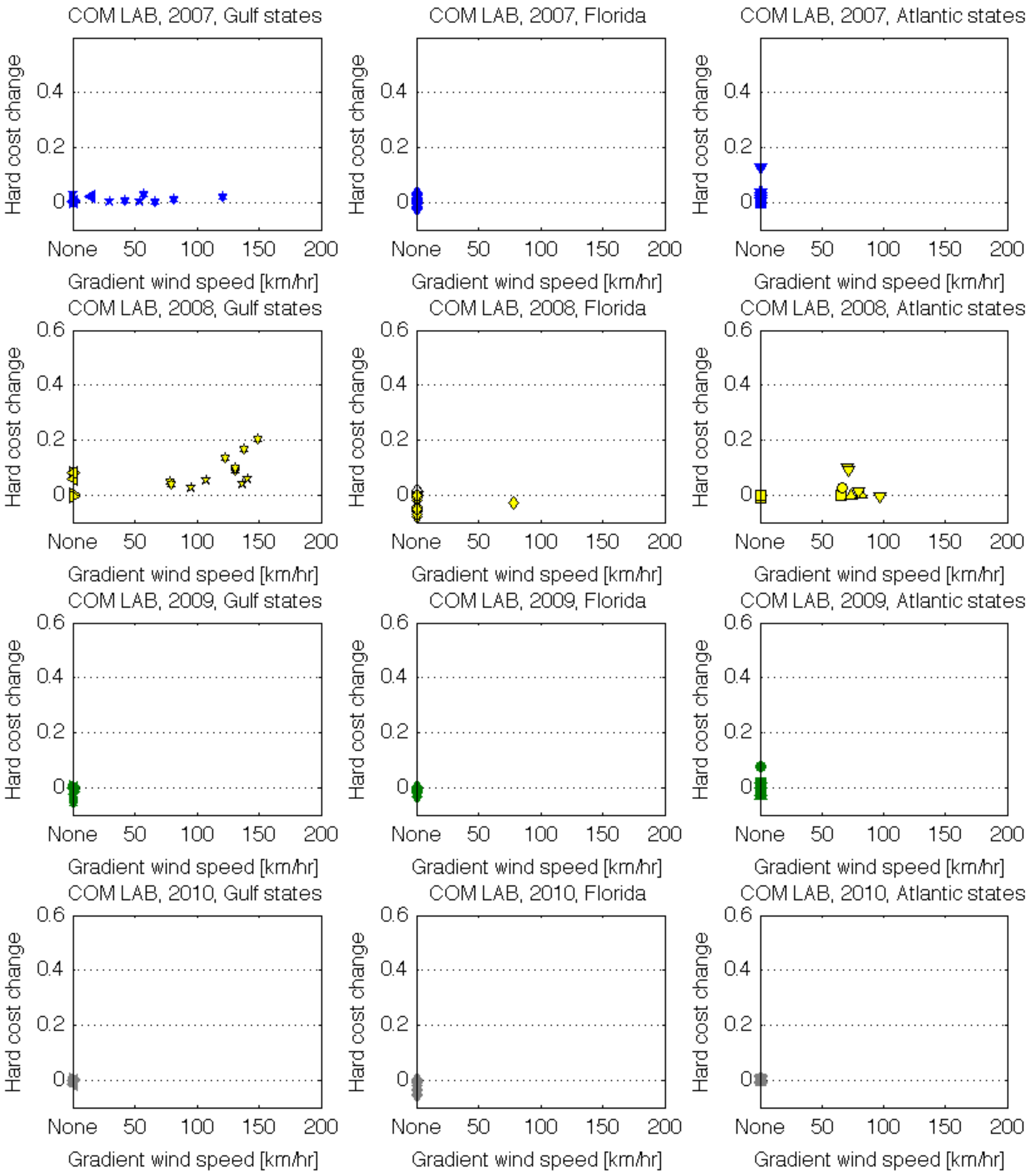

Figure B.11: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the commercial labor component basket.

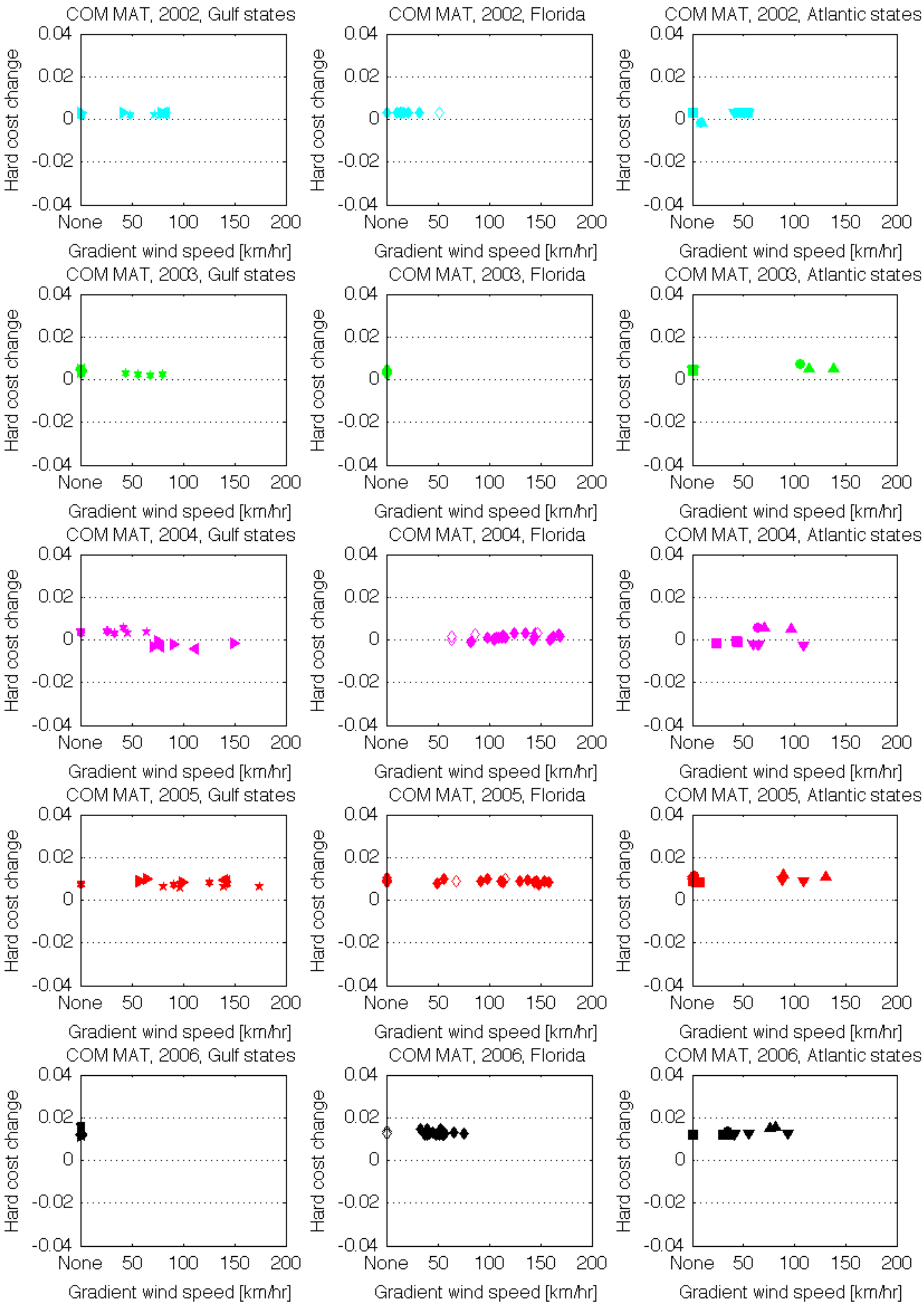

Figure B.12: Observed hard cost change from July to January versus wind speed by U.S. region and year (2002–2006) for the commercial material component basket. Unfilled symbols in the Florida plots (center) are cities in Northern Florida.

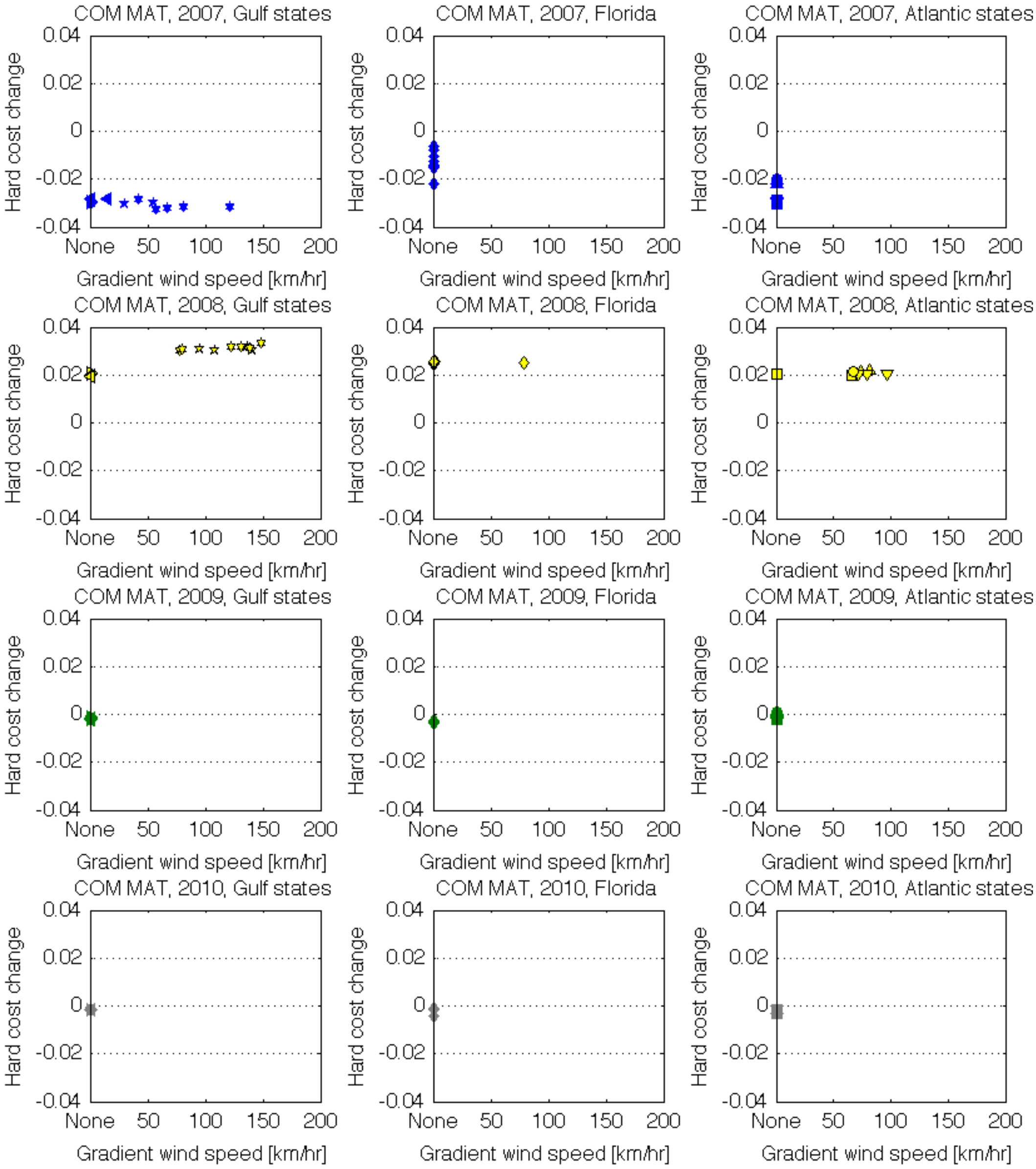

Figure B.13: Observed hard cost change from July to January versus wind speed by U.S. region and year (2007–2010) for the commercial material component basket.

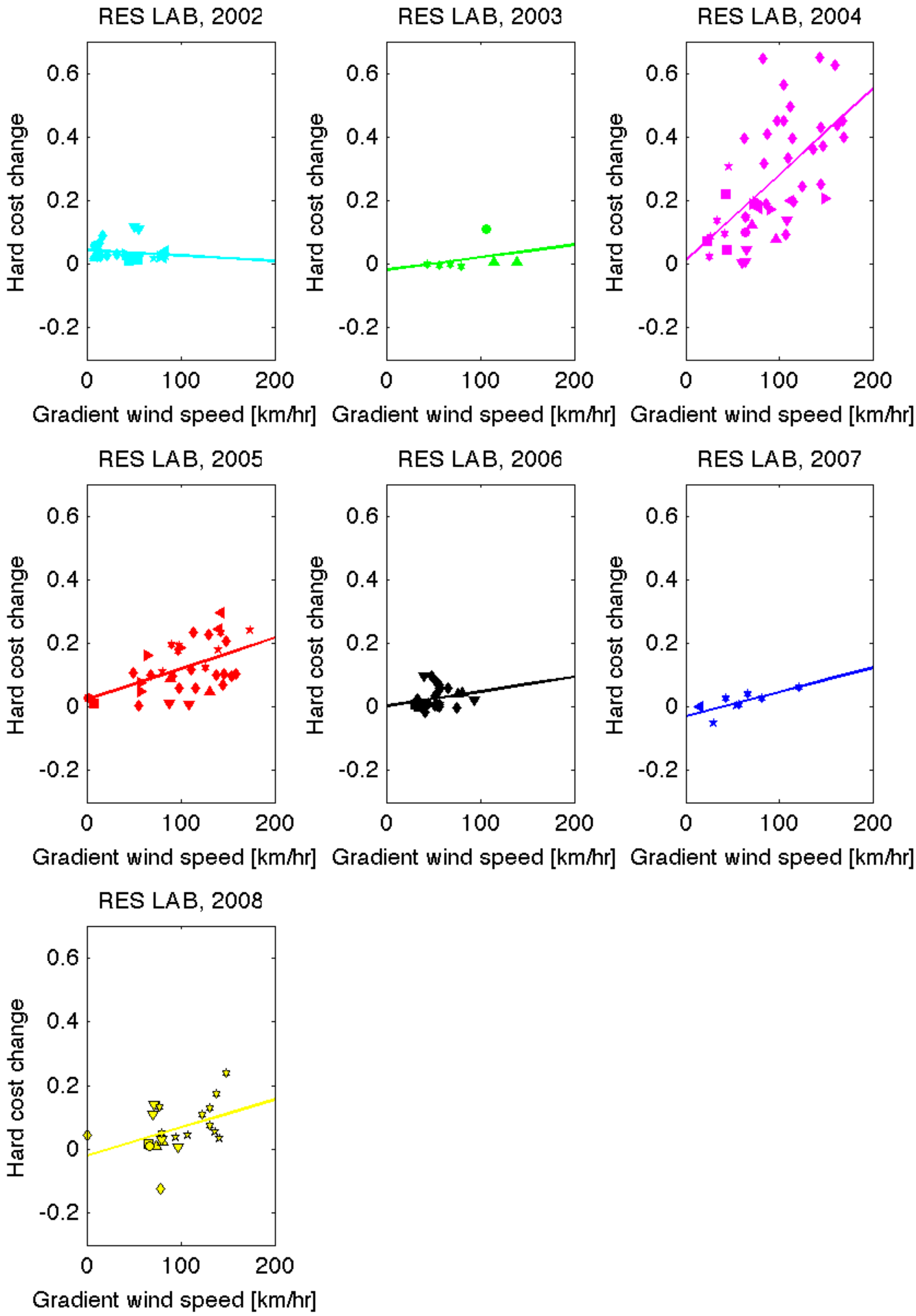

Figure B.14: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the residential labor component basket

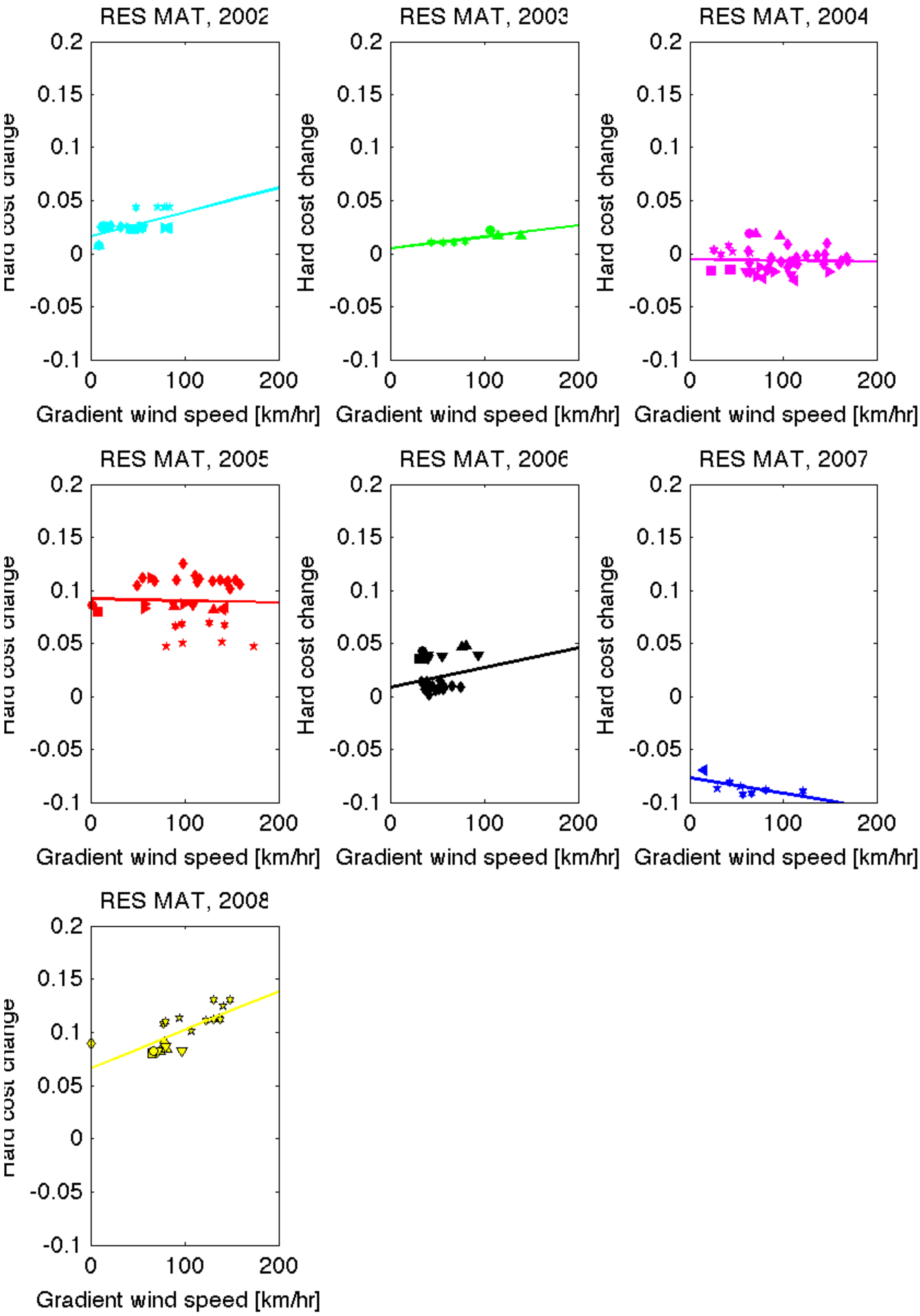

Figure B.15: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the residential material component basket

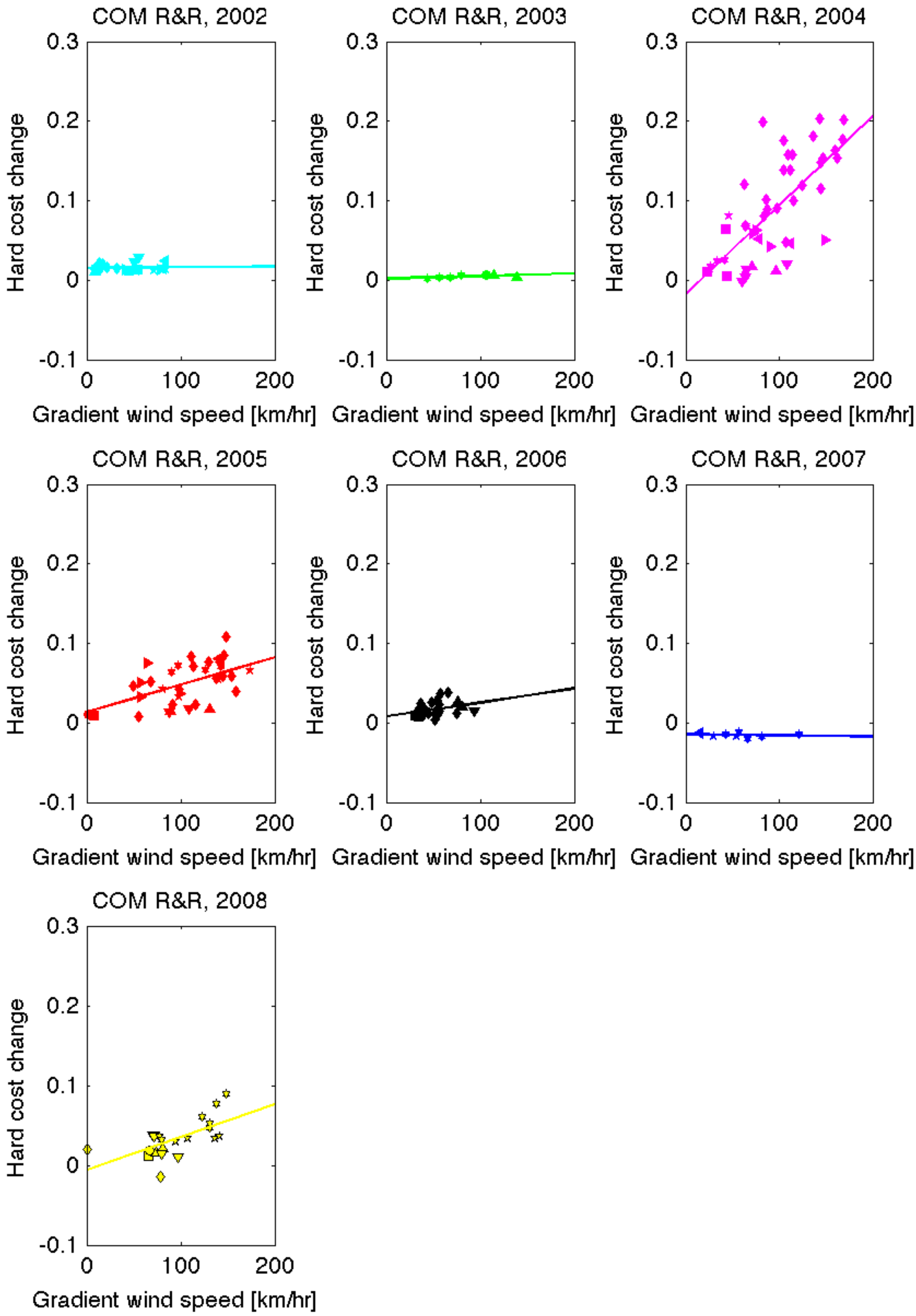

Figure B.16: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the commercial remove & replace basket

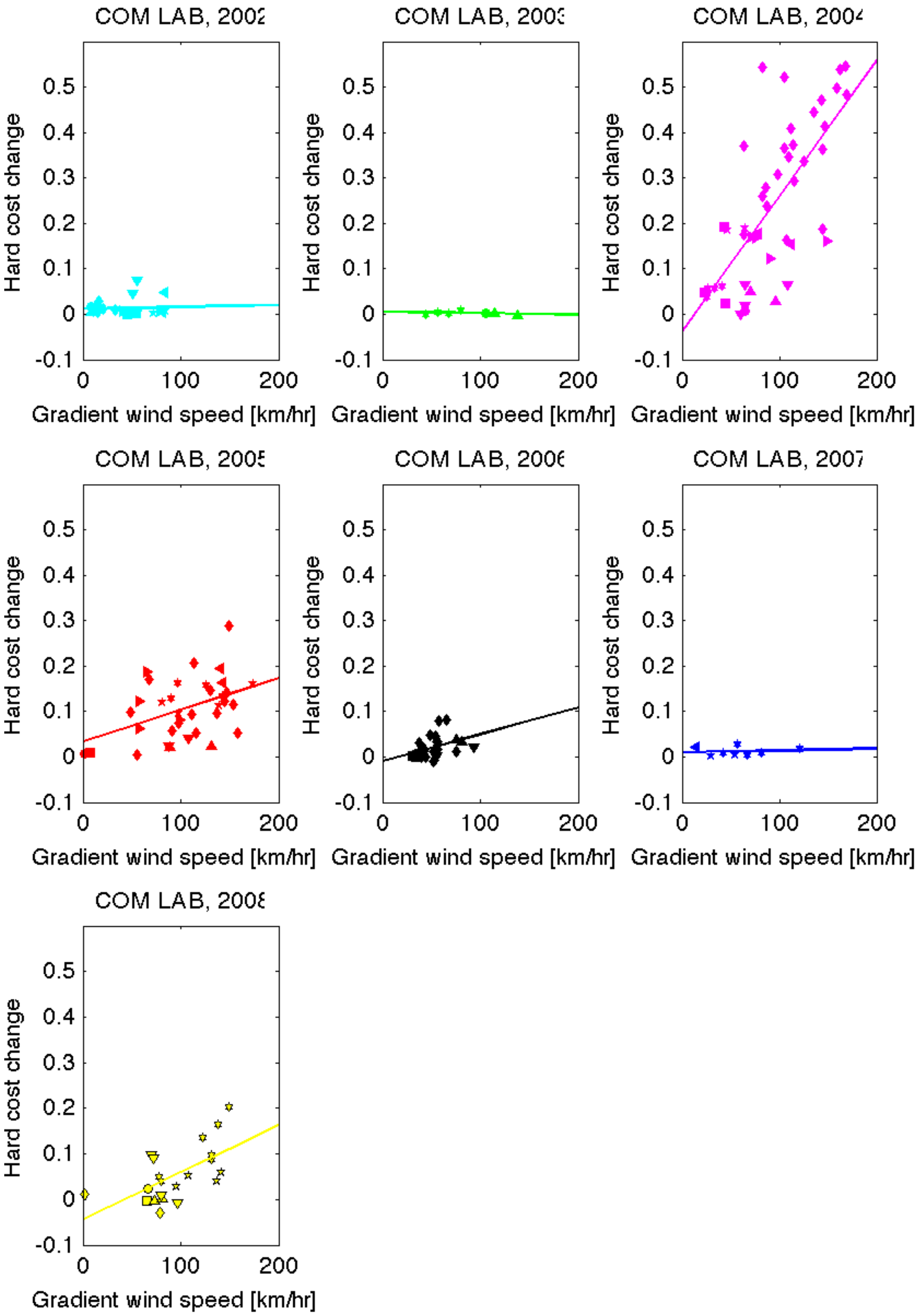

Figure B.17: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the commercial labor component basket

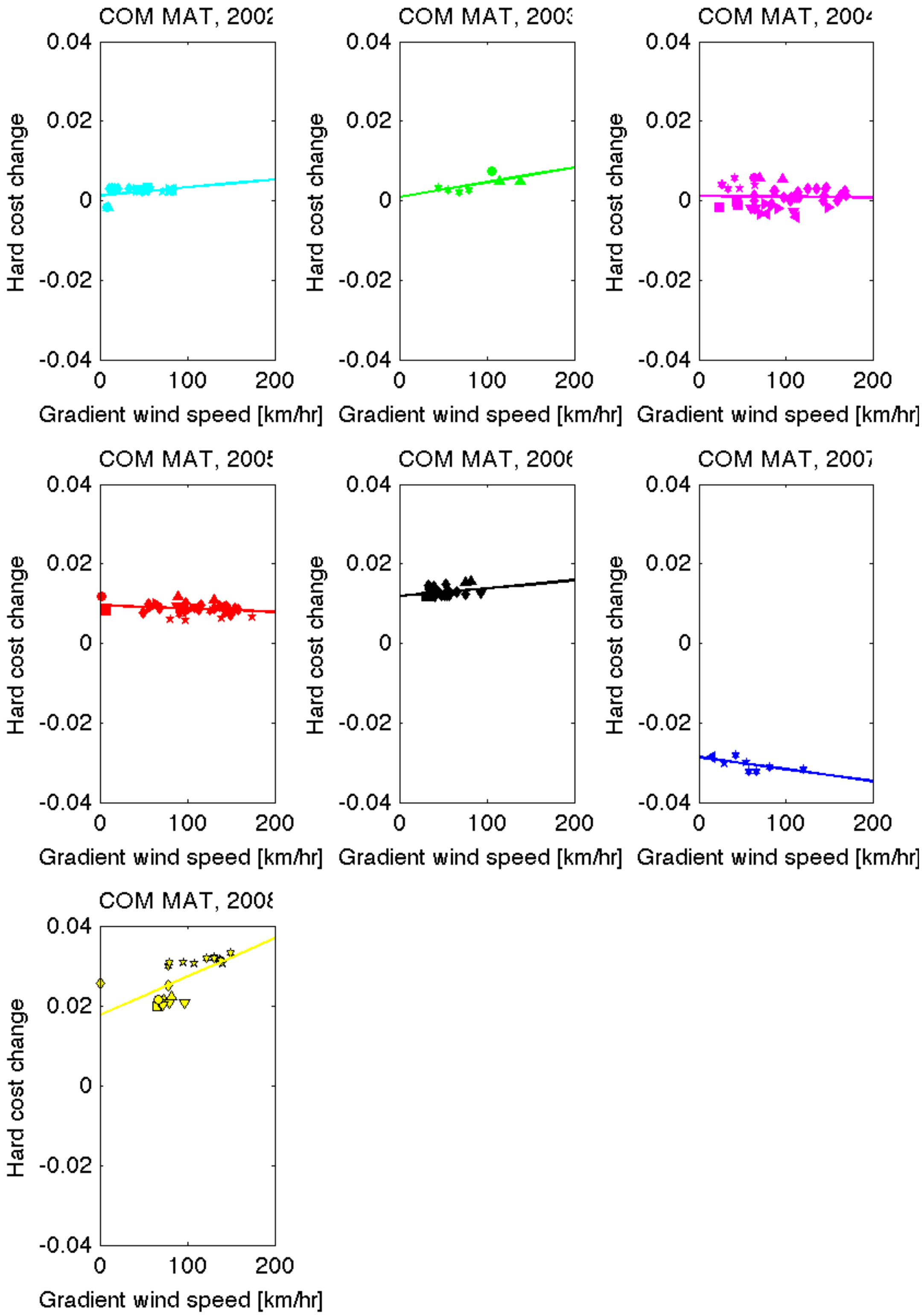

Figure B.18: Observed hard cost change from July to January versus wind speed by year (2002–2008) with simple linear regression for the commercial material component basket

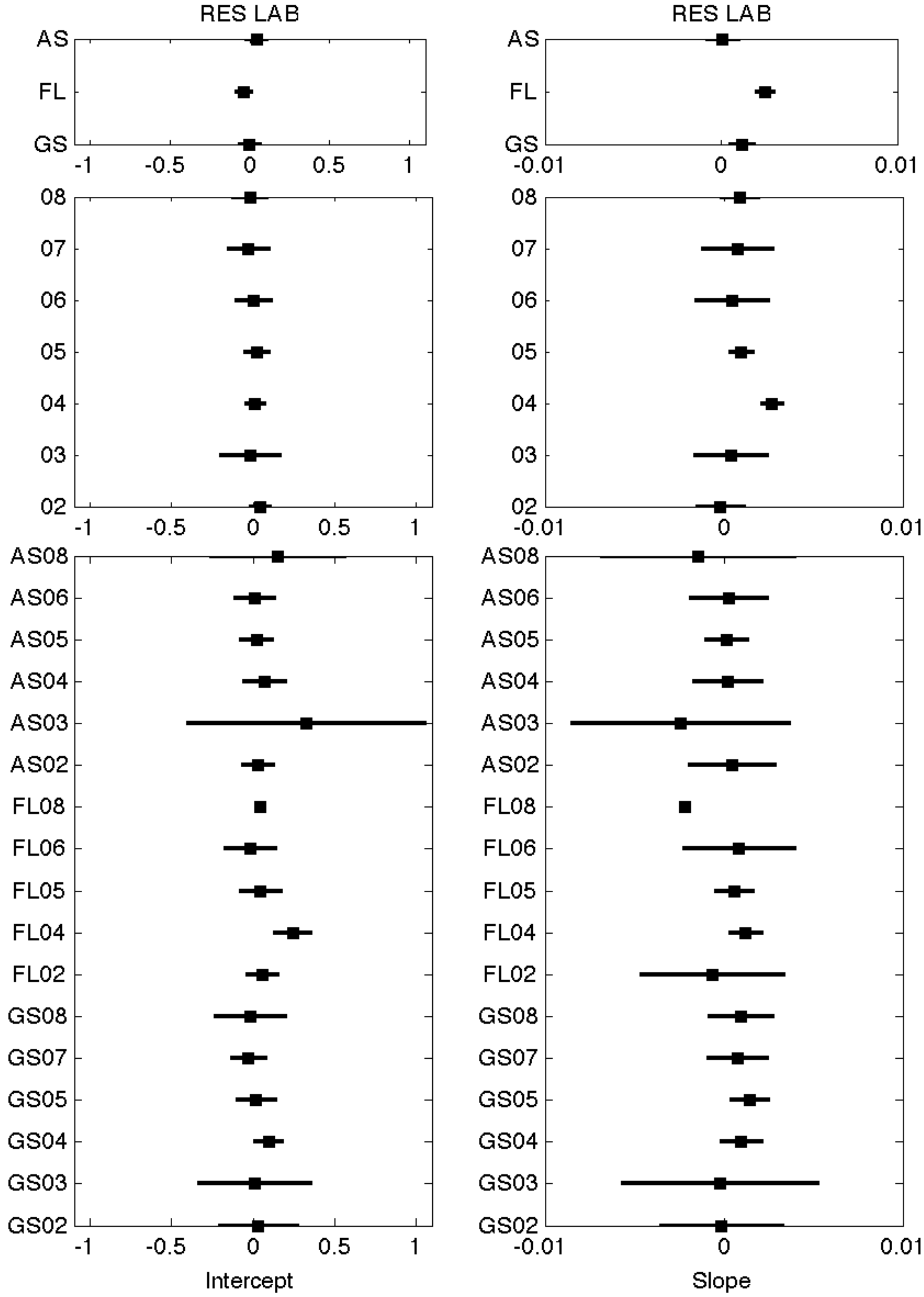

Figure B.19: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the residential labor component basket. “AS” is Atlantic States; “FL” is Florida; and “GS” is Gulf States.

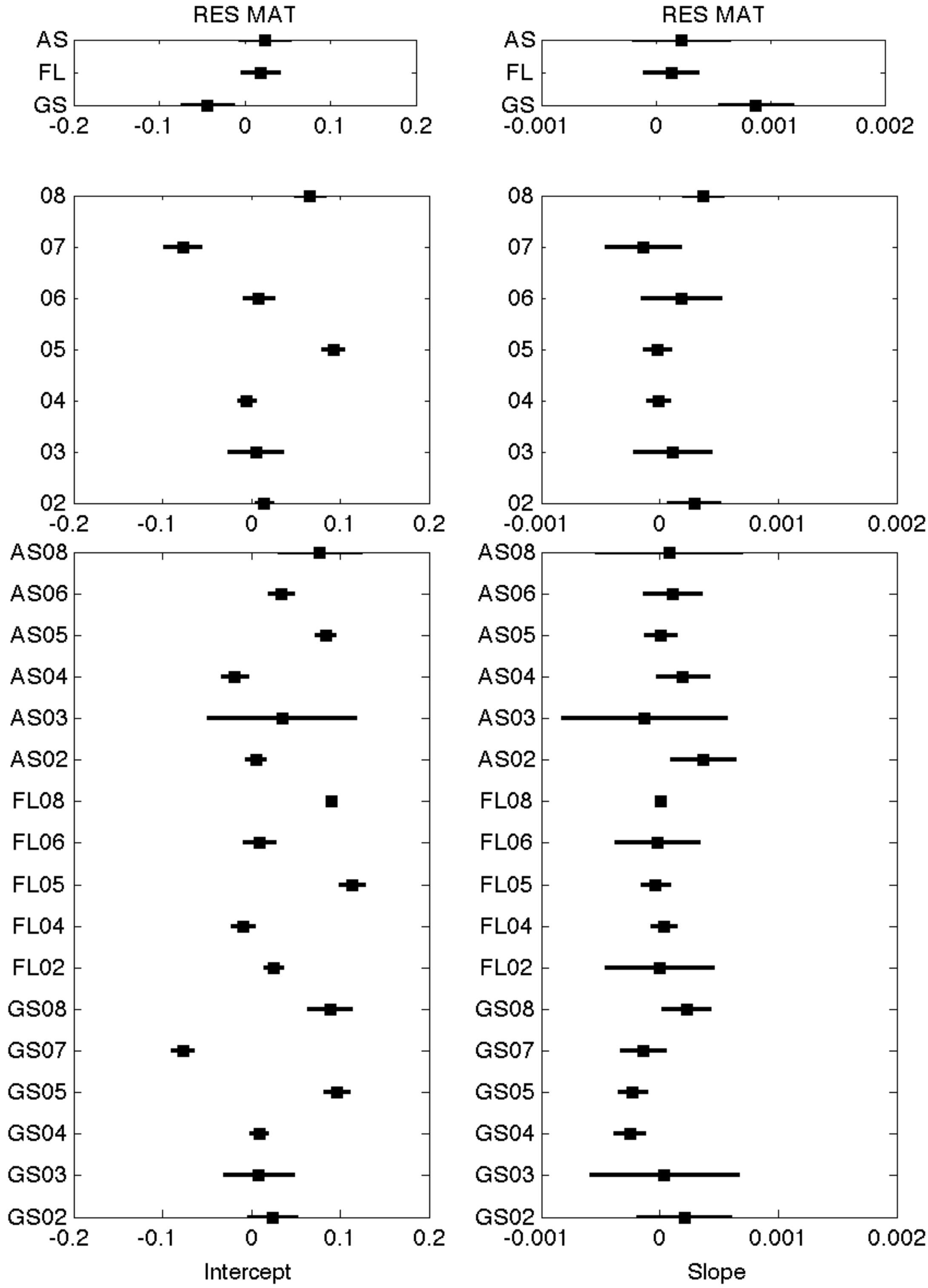

Figure B.20: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the residential material component basket.

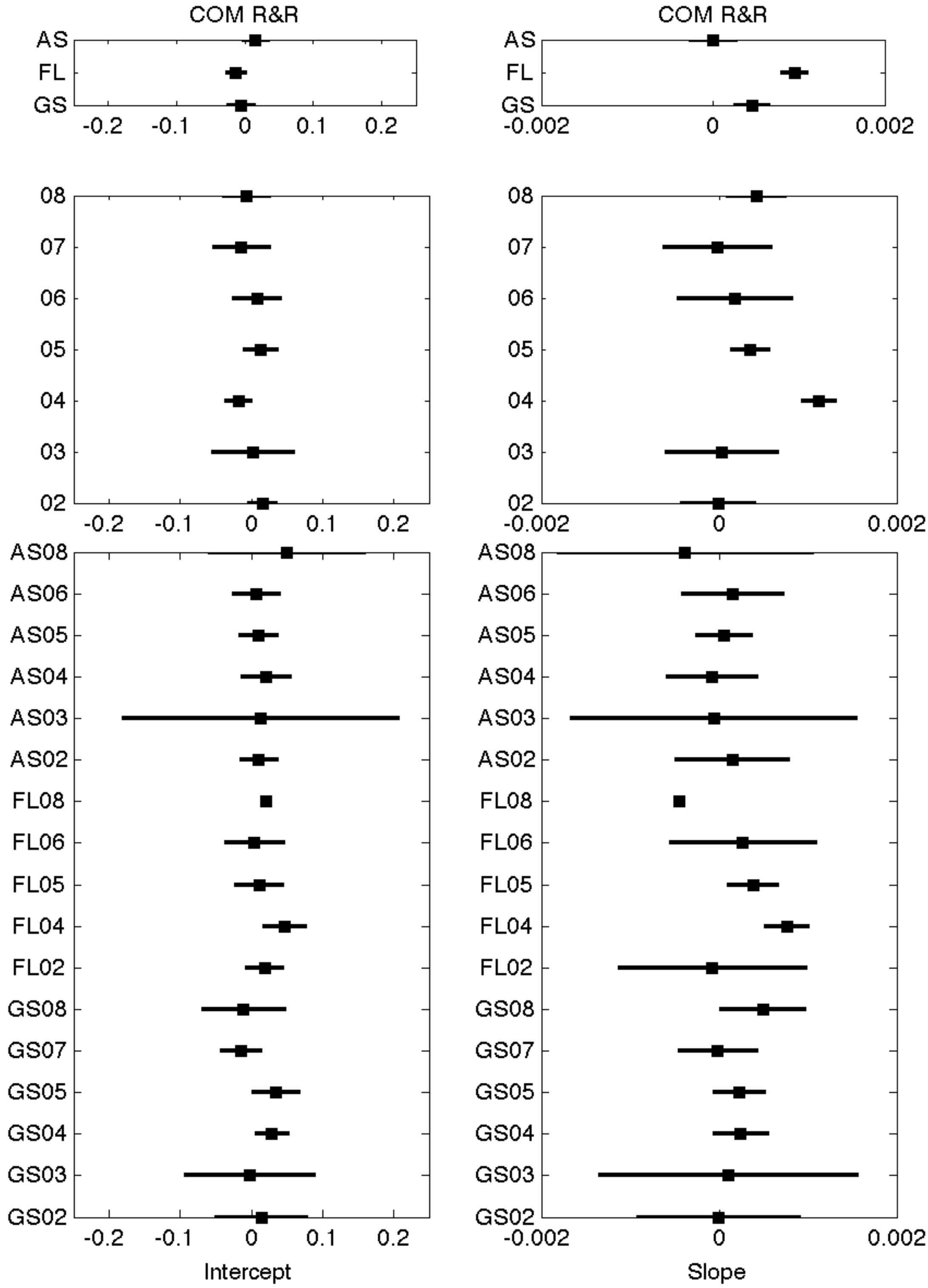

Figure B.21: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the commercial remove & replace basket. "AS" is Atlantic States; "FL" is Florida; and "GS" is Gulf States.

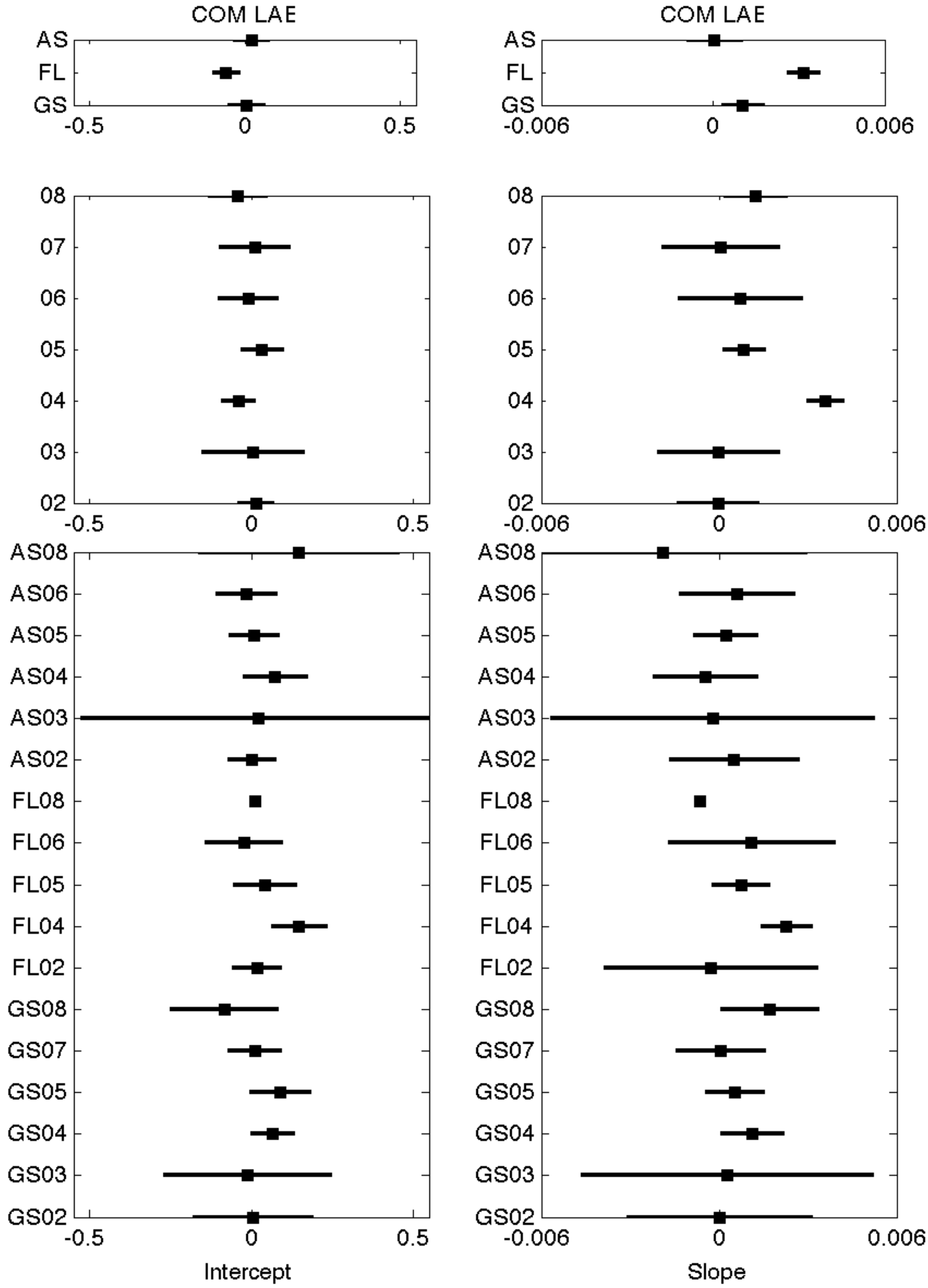

Figure B.22: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the commercial labor component basket.

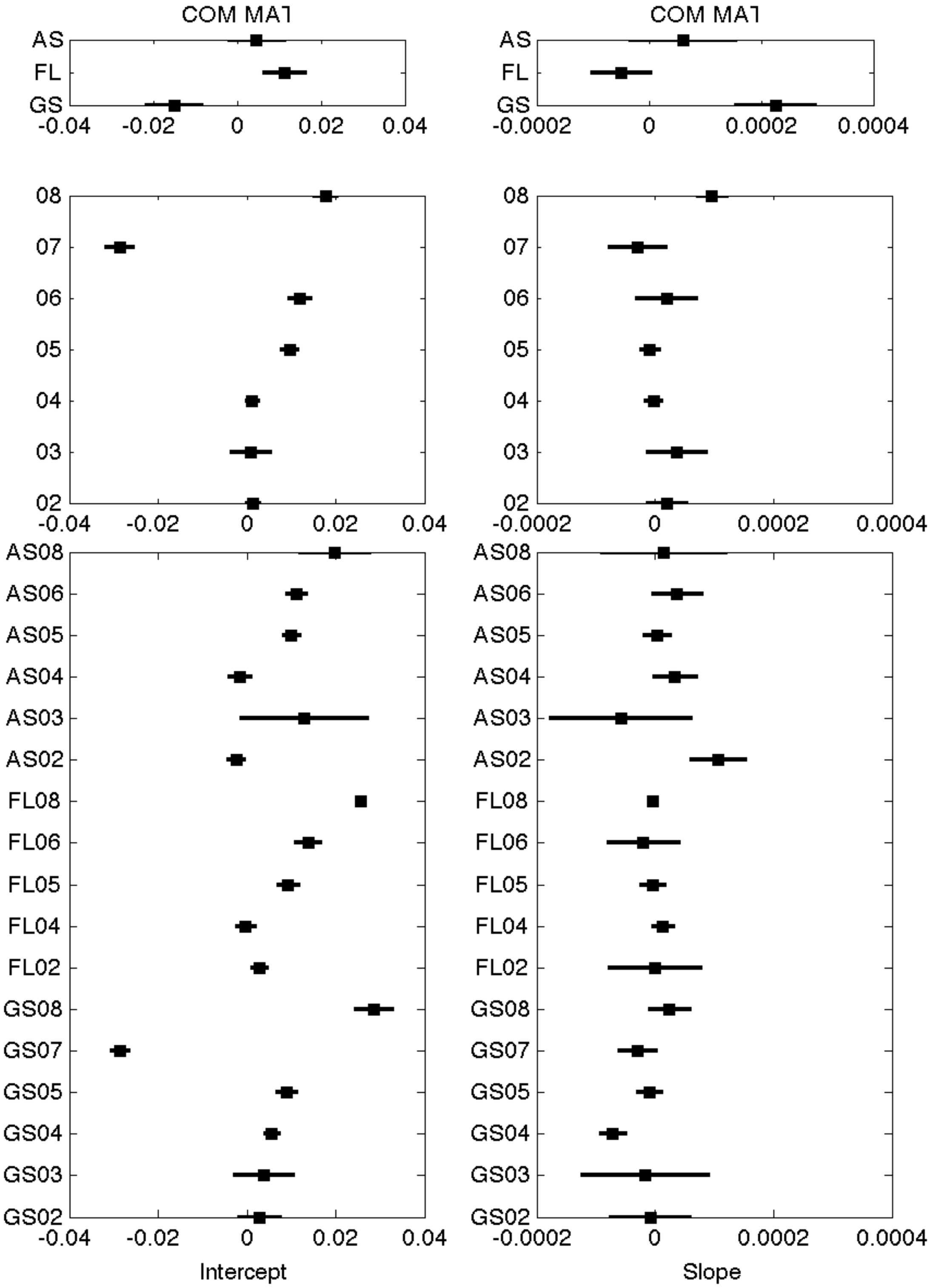

Figure B.23: 95 percent confidence intervals for the intercepts and slopes estimated with simple linear regression by U.S. region (top), by year (middle), or by region and year (bottom), for the commercial material component basket.

City-level model						Log-likelihood	AIC	Δ_{AIC}	weight
	βw_i	$+\gamma_{jk} n_{ci}$		$+\phi_2 \Delta_{2i}$	$+\epsilon_i$	195.1	-380.2	0.0	0.3899
	$\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$		$+\phi_2 \Delta_{2i}$	$+\epsilon_i$	196.9	-379.8	0.4	0.3192
	$\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$			$+\epsilon_i$	195.7	-379.5	0.7	0.2747
	$\beta_{jk} w_i$	$+\gamma n_{ci}$		$+\phi_2 \Delta_{2i}$	$+\epsilon_i$	191.5	-373.1	7.1	0.0112
	$\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1 \Delta_{1i} + \phi_2 \Delta_{2i}$	$+\epsilon_i$	197.8	-371.5	8.7	0.0050
α_{jk}	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	197.8	-359.5	20.7	0.0000

Table B.3: First set of proposed models and their AIC values for hard cost change from July to January of the residential labor component basket

City-level model						Log-likelihood	AIC	Δ_{AIC}	weight
α_{jk}	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	547.3	-1069.0	0.0	0.6263
α_{jk}	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	546.4	-1067.0	2.0	0.2304
α_{jk}	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	549.3	-1065.0	4.0	0.0848
α_{jk}		$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	543.6	-1063.0	6.0	0.0312
α_{jk}	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1 \Delta_{1i} + \phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	548.7	-1061.0	8.0	0.0115
α_{jk}	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	540.6	-1061.0	8.0	0.0115
α_{jk}	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	542.3	-1059.0	10.0	0.0042
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	535.9	-1052	17.0	0.0001
α_{jk}	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	535.1	-1050	19.0	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	531.5	-1047.0	22.0	0.0000
α	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	526.6	-1027.0	42.0	0.0000
α	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	467.8	-915.6	153.4	0.0000
α	$+\beta_{jk} w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	462.3	-904.6	164.4	0.0000
α	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$		$+\epsilon_i$	452.9	-889.8	179.2	0.0000
α	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$		$+\epsilon_i$	403.2	-790.5	278.5	0.0000
α	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2 \Delta_{2i} + \phi_3 \Delta_{1i} \Delta_{2i}$	$+\epsilon_i$	405.0	-790.0	279.0	0.0000
α	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$		$+\epsilon_i$	383.9	-751.8	317.2	0.0000

Table B.4: First set of proposed models and their AIC values for hard cost change from July to January of the residential material component basket

			City-level model			Log-likelihood	AIC	Δ_{AIC}	weight
α	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$			$+\epsilon_i$	417.6	-821.1	0.0	0.5692
α	$+\beta_{jk}w_i$	$+\gamma n_{ci}$			$+\epsilon_i$	414.4	-818.9	2.2	0.1895
α_{jk}	$+\beta_{jk}w_i$	$+\gamma n_{ci}$			$+\epsilon_i$	416.3	-818.6	2.5	0.1631
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$			$+\epsilon_i$	417.7	-815.4	5.7	0.0329
α_{jk}	$+\beta_{jk}w_i$				$+\epsilon_i$	413.3	-814.7	6.4	0.0232
α_{jk}	$+\beta_{jk}w_i$		$+\delta_{jk}w_in_{ci}$		$+\epsilon_i$	416.8	-813.6	7.5	0.0134
α_{jk}	$+\beta_{jk}w_i$			$+\phi_2\Delta_{2i}$	$+\epsilon_i$	413.4	-812.7	8.4	0.0085
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_in_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_i$	417.8	-803.6	17.5	0.0001
α_{jk}	$+\beta w_i$	$+\gamma_{jk}n_{ci}$			$+\epsilon_i$	408.5	-803.0	18.1	0.0001
α	$+\beta w_i$	$+\gamma_{jk}n_{ci}$			$+\epsilon_i$	406.1	-802.1	19.0	0.0000
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_in_{ci}$	$+\phi_1\Delta_{1i}+\phi_2\Delta_{2i}+\phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	417.8	-799.6	21.5	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$			$+\epsilon_i$	403.8	-797.7	23.4	0.0000

Table B.5: First set of proposed models and their AIC values for hard cost change from July to January of the commercial remove & replace basket

			City-level model			Log-likelihood	AIC	Δ_{AIC}	weight
α_{jk}	$+\beta_{jk}w_i$				$+\epsilon_i$	236.2	-460.5	0.0	0.4307
α_{jk}	$+\beta_{jk}w_i$			$+\phi_2\Delta_{2i}$	$+\epsilon_i$	236.5	-459.0	1.5	0.2034
α_{jk}	$+\beta_{jk}w_i$			$+\phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	236.3	-458.5	2.0	0.1584
α_{jk}	$+\beta_{jk}w_i$			$+\phi_1\Delta_{1i}+\phi_2\Delta_{2i}$	$+\epsilon_i$	236.8	-457.5	3.0	0.0961
α	$+\beta_{jk}w_i$				$+\epsilon_i$	232.6	-457.2	3.3	0.0827
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$			$+\epsilon_i$	237.3	-454.6	5.9	0.0225
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$		$+\phi_1\Delta_{1i}+\phi_2\Delta_{2i}$	$+\epsilon_i$	237.9	-451.8	8.7	0.0056
α_{jk}			$+\delta_{jk}w_in_{ci}$		$+\epsilon_i$	229.2	-446.5	14.0	0.0004
α_{jk}	$+\beta w_i$				$+\epsilon_i$	226.0	-443.9	16.6	0.0001

Table B.6: First set of proposed models and their AIC values for hard cost change from July to January of the commercial labor component basket

			City-level model			Log-likelihood	AIC	Δ_{AIC}	weight
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	822.5	-1617.0	0.0	0.4209
α_{jk}	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	826.0	-1616.0	1.0	0.2553
α_{jk}	$+\beta_{jk}w_i$		$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	820.4	-1615.0	2.0	0.1549
α_{jk}	$+\beta_{jk}w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	818.0	-1614.0	3.0	0.0939
α_{jk}	$+\beta_{jk}w_i$	$+\gamma n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	820.6	-1613.0	4.0	0.0570
α_{jk}		$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	817.8	-1610.0	7.0	0.0127
α_{jk}	$+\beta w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	817.9	-1608.0	9.0	0.0047
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	813.0	-1604.0	13.0	0.0006
α	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	810.5	-1593.0	24.0	0.0000
α_{jk}	$+\beta w_i$	$+\gamma_{jk}n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	802.7	-1583.0	34.0	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	798.2	-1578.0	39.0	0.0000
α	$+\beta_{jk}w_i$	$+\gamma_{jk}n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	766.6	-1511.0	106.0	0.0000
α	$+\beta w_i$	$+\gamma_{jk}n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	755.6	-1489.0	128.0	0.0000
α	$+\beta w_i$	$+\gamma_{jk}n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	744.4	-1471.0	146.0	0.0000
α	$+\beta_{jk}w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	680.3	-1343.0	274.0	0.0000
α	$+\beta_{jk}w_i$	$+\gamma n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	680.4	-1339.0	278.0	0.0000
α	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk}w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_i$	670.9	-1324.0	293.0	0.0000

Table B.7: First set of proposed models and their AIC values for hard cost change from July to January of the commercial material component basket

City-level model						*Log-likelihood*	AIC	Δ_{AIC}	*weight*
α	$+\beta_{jk}w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_{jk}$	266.0	-486.1	0.0	0.8311
α	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_{jk}$	264.4	-482.9	3.2	0.1689
α	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_{jk}$	252.8	-459.7	26.4	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i}$	$+\epsilon_{jk}$	252.4	-458.9	27.2	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	252.9	-457.9	28.2	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$		$+\epsilon_{jk}$	250.9	-457.8	28.3	0.0000
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	253.3	-456.7	29.4	0.0000

Table B.8: Second set of proposed models and their AIC values for hard cost change from July to January of the residential labor component basket

City-level model				*Log-likelihood*	AIC	Δ_{AIC}	*weight*
α_{jk}		$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	599.4	-1154.9	0.0	0.3555
α_{jk}			$+\epsilon_{jk}$	596.4	-1154.8	0.1	0.3423
α_{jk}	$+\gamma n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	599.6	-1153.3	1.6	0.1562
α_{jk}	$+\gamma n_{ci}$		$+\epsilon_{jk}$	596.6	-1153.1	1.8	0.1460

Table B.9: Second set of proposed models and their AIC values for hard cost change from July to January of the residential material component basket

City-level model						*Log-likelihood*	AIC	Δ_{AIC}	*weight*
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	513.3	-978.7	0.0	0.5050
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	514.1	-978.1	0.6	0.3835
α	$+\beta w_i$	$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	511.8	-975.7	3.0	0.1115
α	$+\beta_{jk} w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	497.1	-946.2	32.5	0.0000
α	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	493.9	-939.8	38.9	0.0000

Table B.10: Second set of proposed models and their AIC values for hard cost change from July to January of the commercial remove & replace basket

City-level model						*Log-likelihood*	AIC	Δ_{AIC}	*weight*
α_{jk}		$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	325.5	-605.0	0.0	0.4370
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	326.1	-604.2	0.8	0.2939
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	326.3	-602.6	2.4	0.1344
α		$+\gamma_{jk} n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	324.0	-602.0	3.0	0.0972
α_{jk}		$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i}$	$+\epsilon_{jk}$	322.0	-600.1	4.9	0.0375
α		$+\gamma n_{ci}$	$+\delta_{jk} w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	314.9	-583.8	21.2	0.0000

Table B.11: Second set of proposed models and their AIC values for hard cost change from July to January of the commercial labor component basket

City-level model						*Log-likelihood*	AIC	Δ_{AIC}	*weight*
α_{jk}				$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	923.4	-1802.7	0.0	0.4683
α_{jk}		$+\gamma n_{ci}$		$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	923.8	-1801.7	1.1	0.2728
α_{jk}				$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i}$	$+\epsilon_{jk}$	921.6	-1801.2	1.6	0.2149
α_{jk}	$+\beta w_i$	$+\gamma n_{ci}$	$+\delta w_i n_{ci}$	$+\phi_1\Delta_{1i} + \phi_2\Delta_{2i} + \phi_3\Delta_{1i}\Delta_{2i}$	$+\epsilon_{jk}$	924.0	-1798.0	4.7	0.0440

Table B.12: Second set of proposed models and their AIC values for hard cost change from July to January of the commercial material component basket