Dear Mr. Chairman:

Pursuant to your request, the Congressional Budget Office (CBO) has reviewed the study commissioned by the Federal Emergency Management Agency (FEMA), as required by section 578 of the National Flood Insurance Reform Act of 1994, on the effects of eliminating subsidies in the National Flood Insurance Program.

The questions addressed by that study are difficult, and the available data limit the study’s ability to answer them precisely. As discussed in the attachment, CBO believes that the study’s methodology is basically sound; however, its specific numerical findings are accompanied by a great deal of uncertainty and should be interpreted with caution. For example, CBO agrees that eliminating the subsidy immediately would lead a significant share of policyholders to drop their coverage but is not convinced that that share would be close to the study’s estimate of 23 percent. In short, CBO considers the study less valuable for its quantitative results than for its broad qualitative conclusions.

The attachment was written by Perry Beider of CBO’s Microeconomic and Financial Studies Division. CBO appreciates the cooperation provided by staff of FEMA and its contractor, PricewaterhouseCoopers, in conjunction with this review.

I hope the attached information will be useful to you. If you have any questions about it or the FEMA report, please call me, or have your staff contact Perry at 226-2940.

Sincerely,

Dan L. Crippen

Attachment

Identical letters sent to:
Honorable Bob Ney
Honorable Barney Frank
CBO's Review of a Study of the Economic Effects of Charging Actuarially Based Premium Rates for Federal Flood Insurance

September 2000

To satisfy requirements of section 578 of the National Flood Insurance Reform Act of 1994, the Federal Emergency Management Agency (FEMA) contracted with PricewaterhouseCoopers (PwC) to study the economic effects of eliminating the government subsidies in the National Flood Insurance Program. Under current law, flood insurance premiums are generally subsidized for “pre-FIRM structures”—those built before the completion of a participating community’s Flood Insurance Rate Map (or before 1975, whichever is later).

At the request of the Subcommittee on Housing and Community Opportunity of the House Committee on Banking and Financial Services, the Congressional Budget Office (CBO) has reviewed the resulting report, titled Study of the Economic Effects of Charging Actuarially Based Premium Rates for Pre-FIRM Structures. CBO’s analysis included three components:

- Reviewing the report itself;
- Reviewing FEMA’s own internally sponsored review—commissioned from Professor Richard N. Boisvert of Cornell University—of the report’s “draft final” version; and
- Sending three rounds of questions to FEMA and its contractor about assumptions and methods not adequately described or explained in the report.1

CBO appreciates the efforts and open spirit of FEMA staff in cooperating with this review.

CBO concurs with FEMA that the basic methods and assumptions used in the study were reasonable and provide adequate support for its findings—viewed in broad, qualitative terms. However, CBO also concludes that the study’s specific quantitative estimates—particularly for individual communities or groups of communities—are subject to great uncertainty because of important limitations in the

---

1. In addition to its omissions and ambiguities, the report apparently also contains a few errors. The most important one involves the approach used to extrapolate national results from the results for the study’s 50 sample communities. In response to a criticism of the approach described in Section 9 and Appendix E of the report, PwC staff orally assured CBO that the description is incorrect, reflecting early methods that were discarded as additional information was generated during the course of the study. According to PwC staff, the formula actually used to extrapolate the study’s national results is the same one CBO had recommended.
available data. Two of the most critical limitations concern the number of communities and structures directly observed in the study and the extent to which demand for flood insurance is affected by price and other factors.

Thus, for example, CBO finds adequate support for the study's finding that a significant share of policyholders would drop their coverage if the subsidy was eliminated. But it finds scant basis for confidence that the specific share would be the estimated 23 percent, rather than 10 percent or 40 percent. CBO interprets the study’s other major findings in similarly qualitative terms: as showing that eliminating the subsidy would lead to very large percentage increases in insurance premiums for buildings lying six or more feet below the base flood elevation, or BFE (the water height expected during a 100-year flood); that many pre-FIRM structures lie above BFE and therefore would be eligible for lower premiums under post-FIRM rates; and that some communities would experience discernible losses in average property values and property tax collections. One other implication of the results, not highlighted in the study, is that the National Flood Insurance Program (NFIP) involves subsidies and cross-subsidies that are large in percentage terms. (See the appendix on page 12 for a brief discussion of that issue.)

The rest of this review discusses the rationale for CBO’s conclusions. The following sections outline the study’s data and methods, identify its fundamental limitations, and provide some specific interpretations and clarifications about the results for individual communities and groups of communities.

OVERVIEW OF THE STUDY’S DATA AND METHODS

FEMA’s study had ambitious statutory goals, which included estimating the number and types of properties nationwide that would be affected by reducing or eliminating the subsidies and estimating the resulting effects on premiums, participation in the program, property values, and property tax revenues. To meet those goals, PwC undertook extensive field surveys and data collection and created an elaborate simulation model. Information on the study’s methodology is available both in the study itself and in FEMA’s “Recommendations” memo, but a brief overview here is useful in explaining PwC’s approach and highlighting some of the central building blocks of its analysis.

The study focused on 50 sample communities, selected from the more than 15,000 NFIP communities nationwide that have 100-year floodplains (known as special flood hazard areas, or SFHAs). For each of those sample communities, the

2. More than 19,000 communities nationwide participate in the NFIP, but some 1,800 lack an SFHA and hence have no structures insured at subsidized rates. Another 2,000 NFIP communities were omitted
analysis grouped the SFHA structures into “cells” defined by type of occupancy (single-family detached, multi-family attached, high-rise condominium, manufactured housing, or nonresidential), age category (less than 21 years old, 21 to 50, or over 50), flood zone (coastal or inland), FIRM status (pre-FIRM or post-FIRM), presence or absence of a basement, and elevation of the lowest floor relative to the base flood elevation (at or above BFE, 1 to 2 feet below, 3 to 5 feet below, or 6 or more feet below). To illustrate, one such cell comprised pre-FIRM, single-family detached homes more than 50 years old, situated 3 to 5 feet below BFE in an inland flood zone, and lacking a basement. Some parts of the analysis further classified structures by condition (at or above average, or below average) and subdivided single-family detached residences into one-story and multistory structures.

PwC estimated the number and characteristics of structures in each cell and community using existing data, primarily from the Census Bureau, supplemented by direct surveys of 7,628 pre-FIRM structures in 23 of the 50 communities. In particular, PwC was able to estimate the distributions of structures in each community by occupancy type, age, FIRM status, condition, and flood zone on the basis of existing data. But it had to impute the needed information on elevation, basement presence, number of floors (for single-family detached homes), property value, and replacement cost from the observations of the surveyed structures.

The analysis tailored the imputation methods to the nature of the needed data and the available information. For example, because the prevalence of basements varies among regions of the United States, PwC based the percentage incidence of basements in cells of the 27 nonsurveyed communities on weighted averages of the percentages in corresponding cells of the surveyed communities within the same regions. In turn, PwC assigned the cells for a given type of structure in a surveyed community the incidence of basements found among the surveyed structures of that type in that community. As discussed below, imputing property values required a particularly complex chain of methods, in part because direct observations of those values were available for only 2,339 of the 7,628 surveyed structures.

Once the data collection and imputation for the various cells of structures in the 50 communities was complete, PwC simulated the effects of eliminating or phasing out the pre-FIRM premium rates through a series of models that tracked the evolution over time of the premiums, NFIP participation, mortgage status, physical

---

3. PwC assigned structures whose elevations were directly observed into an elevation category by rounding to the nearest integer. Thus, a building whose lowest floor was 0.4 feet below BFE was assigned to the “at or above BFE” category.
condition, and property value of the structures in each cell and community. The models incorporated dozens of assumptions about population growth, property turnover, mortgage refinancing, compliance with mandatory purchase requirements, the random incidence and impact of floods, property deterioration and abandonment, capitalization of insurance prices into property values, and other factors. The majority of the assumptions were either well grounded in available evidence, limited in their impact on the main findings, or both. But the sheer number of assumptions makes them collectively a source of uncertainty and potential error in those findings. Moreover, as discussed in the next section, the assumptions about the influences on NFIP participation—particularly the impact of premium rates—were both based on limited evidence and critical to the study’s results.

The last step in the analysis—extrapolating national estimates from the results calculated for the 50 sample communities—also used a cell approach, although the study refers to them as “clusters.” The approach classified each of the study's 15,000 communities in a cell according to its location (coastal or inland), its rate of population growth from 1990 to 1995 (above or below the median), and its number of insured pre-FIRM structures (assigned to one of five “tiers”). Individual cells contained as many as 14 of the 50 sample communities; however, eight of the cells in the larger tiers included just one sample community each. PwC derived the national estimates by assuming that within a given cell, the averages estimated for the pre-FIRM structures in the sample communities applied to the pre-FIRM structures in all such communities nationwide (and similarly for post-FIRM structures). In other words, each community’s results were scaled up according to the number of nationwide structures in similar communities represented by each local pre-FIRM or post-FIRM structure.

LIMITATIONS OF THE STUDY

By definition, a model is a simplification of reality. Accordingly, any model’s projections of the future are accompanied by some degree of uncertainty. In broad terms, the amount of uncertainty is determined by how many of the relevant aspects of reality the model includes and how accurately it represents them.

The study FEMA commissioned from PricewaterhouseCoopers attempted to model a large number of factors relevant to the effects of eliminating the subsidies for pre-FIRM structures, and it involved extensive efforts to collect accurate and

4. The five tiers were 0 to 134, 135 to 555, 556 to 2,449, 2,450 to 9,394, and more than 9,394.

5. The 50 sample communities represented only 16 of the 20 cells in the classification scheme; none were inland high-growth or low-growth communities in the largest tier or coastal low-growth communities in the smallest two tiers. Presumably, few if any NFIP communities nationwide are in those four cells.
representative data. Nonetheless, CBO believes that the information used to conduct the analysis is limited in important ways and, consequently, that a significant degree of uncertainty surrounds its results. The range of uncertainty is unknown: given the complexity of the analysis, PwC was unable to calculate confidence intervals around the study’s findings, and the study does not include any sensitivity results showing the effects of alternative assumptions.\(^6\)

### Sample Size

One obvious source of uncertainty is the limited samples of communities and structures used in the study. On the positive side, the 50 sample communities contain about 5 percent of all pre-FIRM SFHA structures nationwide, PwC staff report, because of the overrepresentation of large communities. And although a much smaller fraction of structures (about 0.2 percent) was surveyed for detailed data on elevation, basement presence, number of floors, and replacement cost, the absolute number of structures (7,628) is comparable to sample sizes regularly used in public opinion polls that yield confidence intervals of a few percentage points.

On the negative side, the aggregate samples do not tell the whole story. Because structures in different types of communities may vary in their characteristics, and the structures’ owners may therefore respond differently to changes in NFIP premiums, the samples of structures and communities contained in each cell of communities are also important. For the cells of smaller communities, the numbers of structures are probably the more important issue; for example, although inland, low-growth, "tier 1" communities were well represented in the sample (accounting for 12 of the 50 communities), PwC surveyed only 82 structures from Niagara, New York, in that cell. Conversely, for the cells of larger communities, eight of which have only a single representative in the sample, the number of sample communities is probably more critical. For instance, the study’s national results might reflect too few nonresidential structures in coastal, high-growth, tier 5 communities if Ft. Lauderdale has proportionately fewer such structures than the average for Miami and other communities in that cell.

---

6. The study briefly discusses one set of alternative assumptions: using 10-year or 100-year floods instead of 25-year floods to simulate the attrition of pre-FIRM structures from flood damage. Appendix C of the study contains tables showing damage rates for all three frequencies of floods, but only the “base case” model of 25-year floods is used elsewhere in the analysis. The alternatives would probably have made little or no difference to the results: modeling 10-year floods instead would have meant selecting more communities to experience random flooding but at a lower rate of flood damage, and the reverse for 100-year floods.
Demand for Flood Insurance

Probably the most important source of uncertainty about the study’s results is not its sample sizes but its limited base of information on how price and other factors influence the demand for flood insurance. Clearly, the effects of changing the NFIP premiums would depend heavily on buyers’ sensitivity to price. Yet the only empirical support for the study’s modeling of price sensitivity comes from a single General Accounting Office (GAO) report from 1983.7 On the basis of monthly data from 1978 through August 1982 on the number of NFIP policies nationwide, GAO estimated a price elasticity of -0.38, implying that each 10 percent increase in flood insurance premiums would lead to a 3.8 percent decrease in policies sold.8

The relevance of the data underlying that GAO estimate of price elasticity is questionable not only because the data are old but also because they reflect prices that are much lower than—in some cases, as little as one-fortieth of—prices that buyers would face under actuarial rates. Any price elasticity above (that is, less negative than) -1.0 implies that the item in question consumes a larger portion of total spending as price rises, reaching and then exceeding the entire spending budget as some threshold price is crossed. Thus, one expects to see price elasticities of -1.0 and below for items with "high" prices.

Recognizing that GAO's elasticity estimate might not be applicable for the range of prices projected in the study, PwC assumed a more elaborate elasticity function, as follows: -0.38 for price increases up to 112 percent (the maximum increase observed in the GAO data), -0.76 (twice the GAO estimate) for increases of 200 percent or more, and values linearly interpolated between -0.38 and -0.76 for increases between 112 percent and 200 percent (see Figure 1). According to PwC staff, however, the analysis did not use that function for all (actual and potential) purchasers of flood insurance: buyers subject to the mandatory purchase requirements (MPR) who pay or would pay their premiums via escrow reportedly were assumed to have elasticities one-half of those values, on the grounds that they could not drop their coverage as readily when prices rise. Thus, such purchasers respond to price increases according to elasticities ranging from -0.19 to -0.38.

In broad, qualitative terms, PwC’s elasticity function is a plausible extension of the GAO estimate. However, its specific quantitative features are not supported

---


8. Actually, GAO estimated the elasticity as -0.38 using a linear specification of demand and as -0.39 using a logarithmic specification; ibid., p. 29. PwC adopted the former for the FEMA study.
PwC staff report that tests conducted during the course of the study showed that the results were not sensitive to reasonable changes in the elasticity assumption. Papers documenting those sensitivity tests are no longer available, however. CBO believes that the elasticity is so central to the analysis that significant changes could not help but have a significant impact on the results. PwC’s tests may have interpreted the “reasonable” range of elasticities more narrowly than is warranted by the meager empirical evidence.

by any empirical evidence, and a wide variety of other, equally plausible assumptions could have produced significantly different study results. The estimates of reductions in NFIP participation would have been even larger under some alternatives (for example, if the elasticity for the largest price increases had been set at -1.0 instead of -0.76) and smaller under others (such as one that maintained the -0.38 figure for price increases up to 300 percent). Similarly, larger “discounts” off the standard elasticity for policyholders with escrow accounts would have yielded smaller reductions in participation, and vice versa.

Judging by the comments in his review of the draft report for FEMA, Professor Boisvert shares CBO’s view of the uncertainty surrounding the elasticity assumptions and their importance to the study’s results. He wrote,

9. **PwC staff report** that tests conducted during the course of the study showed that the results were not sensitive to reasonable changes in the elasticity assumption. Papers documenting those sensitivity tests are no longer available, however. **CBO believes** that the elasticity is so central to the analysis that significant changes could not help but have a significant impact on the results. **PwC’s tests** may have interpreted the “reasonable” range of elasticities more narrowly than is warranted by the meager empirical evidence.
It is truly unfortunate that this project didn’t include a specific component to collect sufficient data to conduct a credible analysis of the demand for flood insurance. . . . Without more confidence in the estimates of the price elasticity of demand for flood insurance, the value of the resources spent in the current study to collect more accurate elevations and other physical information for pre-FIRM structures in a representative sample of communities is certainly diminished. The validity of the final results is only as good as the accuracy of the individual components of the model. None is more critical to this study than the price elasticity of the demand for insurance. As near as I can tell, no effort was made to develop a reliable estimate.

The assumptions about price elasticity were not the only questionable ones in the model of demand for flood insurance. According to PwC staff, although the model viewed policyholders who are subject to the MPR and pay via escrow as less sensitive to price changes, it considered them no less likely than other policyholders to drop their coverage from one year to the next in the absence of price changes (that is, it assigned them the same annual retention rate) and no more likely to comply with the MPR when they first take out a mortgage that triggers the purchase requirement. In contrast, the modeled rates of policyholders’ retention and initial MPR compliance did vary according to the elevation category of their structures, on the plausible grounds that owners of structures at greater risk (or their lenders) will be more vigilant about buying and keeping flood coverage, all other things (such as premium rates) being equal. Why elevation but not escrow status should affect retention and initial MPR compliance, while escrow status did affect price sensitivity, is unclear. Those assumptions about the influence or lack of influence of escrow payment could affect the model’s results if different states or localities vary in their use of escrow accounts or if escrow use is correlated with other relevant variables.

Data Imputation

Another notable source of uncertainty is the number of assumptions required to complete the full set of cell-level data on structures in each of the 50 sample communities. As discussed above, PwC had to impute data on elevation, basement presence, number of floors, property value, and replacement cost from observations of the surveyed structures.10 Fundamentally, imputations assume that certain

---

10. Also, in estimating the distributions of structures in each community by type, age, FIRM status, condition, and flood zone, PwC had to make some assumptions to supplement the census and FEMA data. For example, inferring the numbers of nonresidential structures required a chain of assumptions involving SFHA population data, county-level data on population and commercial workers, regional Department of Energy data on commercial workers per commercial building, and an adjustment factor.
observed data are good proxies—either directly or after adjustment by averaging, linear regression, or some other technique—for other, unobserved data. As plausible as such an assumption may be, and as efficiently as the imputations may use the information available, they can only approximate the true, missing data. Ideally, the errors in the various approximations will tend to balance out, leaving the ultimate results largely unaffected. But such balance is not guaranteed. Thus, each imputation in a study is a source of uncertainty and possible error in the study's results.

Perhaps the largest uncertainties that result from imputations in the PwC study are those associated with the cell-level property values. Imputing those values was particularly complex because direct observations of property values and structure ages were available for only 2,339 and 2,330, respectively, of the 7,628 surveyed structures. Consequently, the imputation required multiple steps.

First, PwC analysts used the 2,339 known property values to impute values for the other 5,289 surveyed structures. They did so by using the observed values to estimate the coefficients of linear regressions (one for each structure type) and then using the estimated regression functions to predict the missing values. The regressions modeled the property values as functions of the structures’ own replacement costs and county-level population growth, construction cost, and median property value.

Second, PwC used the observed age data to impute age categories (20 to 50 years old or more than 50 years old) for 4,863 of the surveyed structures lacking direct observations. The method for that imputation assumed that structures with sufficiently high property values relative to the census average for that community and type of structure were in the younger category; “sufficiently high” meant above the threshold property-value ratios (again, one for each structure type) chosen to minimize the errors in classifying the structures with known ages.

Third, PwC analysts used their direct and imputed data on the surveyed structures to derive cell-level property values for the 23 surveyed communities. If some of the structures in a given cell in a particular community had been surveyed directly, the analysts used the average of the property values (observed or imputed) for those specific structures as the cell value. For cells that lacked surveyed structures, analysts imputed property values by estimating a log-linear regression for each structure type, specifying property value as a function of elevation category, age

for noncommercial, nonresidential structures such as industrial plants and places of worship.

11. Few of those structures were thought to be in the less-than-20-years-old category.

12. For unspecified reasons, no age category was imputed for 435 of the structures lacking direct observations. Thus, the third step used data on a total of 7,168 structures.
category, flood zone, basement presence, and community dummy variables (to represent many other factors that influence the relative value of a type of structure in a given community). In effect, the regressions allowed PwC to “fill in the blanks” by using average patterns across all 23 surveyed communities.

The final step imputed cell-level property values for the 27 nonsurveyed communities on the basis of those for the surveyed communities. For that imputation, PwC used census data on average property values for each type of structure in each community, calculated the ratios of the cell-level values (obtained in the previous step) to those type-level averages for the surveyed communities, and then applied the ratios to the corresponding type-level averages for the other 27 communities. In particular, the comparisons were conducted within “clusters” of communities with similar type-level average property values, and each ratio was defined as the simple (unweighted) average across a cluster’s surveyed communities of the cell-level property values divided by the simple average of their type-level values.

That elaborate, four-step analysis links many assumptions and approximations to each other. For example, the first step estimates structures’ property values only on the basis of their replacement costs and county-level data, omitting location, lot size, and other specific factors for which data were not available. In turn, those estimated values were used as inputs in both the second step, which assumed that relative property value is a good proxy for structure age, and the third step, which assumed that the explanatory characteristics used in the regressions have roughly the same percentage impact on the values of structures of a given type in all 23 surveyed communities.

Information provided by PwC illustrates the imprecision in the various assumptions. In particular, the threshold ratios chosen in the second step to minimize the misclassification of structures with known ages still yielded error rates of 18 percent or more for four of the five types of structures. That result is not particularly surprising, especially since 60 percent of the property values used to determine the thresholds were themselves imputed in the first step, and those imputed values differ from others in the same community only by virtue of the structures’ replacement costs, which have no logical connection to their ages. The implication, however, is

---

13. The equation shown on page D-13 of the study also includes FIRM status (pre-FIRM versus post-FIRM) as one of the explanatory variables used in the regressions. That inclusion seems to be a mistake: since only pre-FIRM structures were surveyed, the data provided no basis for estimating the impact of FIRM status on property value.

14. If no structures of a given occupancy type were surveyed in a particular community, the regression for that type could not estimate the coefficient for the community’s dummy variable and hence could not suggest the property values for the cells of that type and community. PwC used an alternative method based on census data in such cases.
that the 4,863 imputed ages presumably suffer similar error rates and, hence, that some 10 percent or 15 percent of the age data used in the regressions in the third step are probably wrong. Of course, errors in the first step’s estimates of property values would also affect those regressions.

The point is not that any of PwC’s assumptions or approximations are unreasonable; rather, that they collectively reduce the precision, and perhaps the accuracy, of the study’s findings to some unknown degree.

INTERPRETING THE SUBNATIONAL RESULTS

In light of the methodological limitations discussed above, CBO concludes that the study’s findings should be interpreted primarily in qualitative terms, and the specific numerical results should be treated with caution. That conclusion applies particularly to the results for individual communities and small groups of communities. The estimates for individual communities, shown in Section 8 of the full report, do not allow for as much balancing out of offsetting errors as the national results and therefore may be less reliable. The estimates for the 23 communities with surveyed structures may be somewhat more reliable than those for the 27 nonsurveyed communities, but even the former estimates rest on many assumptions that may be more accurate at the national level.

The report clusters the sample communities into five groups in discussing the range of possible community-level economic impacts. The groups include between six and 14 communities each, as shown in Table ES.2 in the executive summary. Again, the results for those groups of communities should be interpreted with particular care. Although the groupings do allow for some cross-community balancing of errors, the numbers of communities shown in Table ES.2 provide no meaningful information on the relative sizes of the groups nationwide, because the 50 communities were not selected in a purely random sample.

FEMA’s summarizations of the results on community-level effects also warrant some clarification. In the section of the agency’s “Recommendations” memo on “Results of Immediate Subsidy Elimination,” the third bullet says that “Communities with a high portion of properties in the SFHA that are significantly below the BFE could expect property value declines of around 10 percent.” The fifth bullet reports that such communities would face reductions of 14 percent in local property tax revenues. However, those figures, and all of the report’s estimates of effects on property values and property tax revenues, refer only to properties within the SFHA. Depending on how much of a community’s total land area—or more particularly, its total property value—lies within the SFHA, the percentage impact
on the community as a whole could be significantly smaller than the impact on its SFHA properties.

APPENDIX: ESTIMATES OF THE NFIP SUBSIDIES

The estimates provided in the study can be used to shed light on the size of the internal and external subsidies currently used to support the National Flood Insurance Program. Although the specific numerical estimates are subject to the uncertainties discussed earlier, they illustrate the qualitative point that the subsidies and cross-subsidies appear to be large in percentage terms. For example, the estimates imply that roughly one-quarter of all premium dollars from policyholders paying pre-FIRM rates represent an internal cross-subsidy from structures at or above the base flood elevation to structures below BFE. That share follows from the study’s estimates that 2 million of the 4.3 million pre-FIRM structures (44 percent of the total) are at or above BFE and would qualify for lower premiums—averaging just $260, instead of $580 under post-FIRM rates if their elevations were certified.

Moreover, even with that cross-subsidy, pre-FIRM policyholders as a group pay only 40 percent of the average actuarial premium, according to the study’s estimates. Consequently, the NFIP as a whole and the nation’s taxpayers bear an estimated implicit subsidy of roughly $900 per year for each pre-FIRM policy. Like any actuarial estimate of disaster costs, that figure is a long-run average covering many years in which the realized cost of the subsidy is much smaller, or even negative, and a few years in which it is many times larger.