

Incorrect Data Analysis Used To Discredit Seller Funded Down Payment Assistance

A recent letter from Senator Diane Feinstein to a constituent, who had asked the Senator to support down payment assistance, is an illustration of how the Senate has been influenced by the misleading data analysis HUD has put forth about Seller Funded Down Payment Assistance (SFDPA). In her letter, the Senator states:

Foreclosure rates for seller-funded down payment assistance loans have been found to be **three times higher than other FHA loans**.

This oft-quoted statistic, highlighted above, was put forth by HUD in their rationalization of a proposed rule banning down payment assistance.¹ HUD stated that “to-date claim rates ... exceed three times those of borrower-funded purchase loans. ...” and cited the data in Table 5 of that proposed rule to support this statement. Table 5 is produced below. The conclusions HUD draws from Table 5 suffer from the same defect all of HUD’s conclusions suffer from— HUD assumes that if two things are somehow correlated, that one is causing the other.

TABLE 5.—DATE CLAIM RATE COMPARISONS ON FHA-INSURED HOME PURCHASE LOANS BY SOURCE OF DOWNPAYMENT FUNDS AND FISCAL YEAR OF INSURANCE ENDORSEMENT

Fiscal year of insurance endorsement	To-date claim rates in percent					Ratios to “borrower” to-date claim rates			
	Borrower	Family	Nonprofit	Govt agency	Employer	Family	Nonprofit	Govt agency	Employer
2000	6.29	8.38	16.07	13.58	9.52	1.33	2.56	2.16	1.51
2001	5.67	6.68	16.23	13.34	7.24	1.18	2.86	2.35	1.28
2002	4.45	4.58	13.27	10.72	6.16	1.03	2.98	2.41	1.38
2003	3.31	3.58	11.22	8.84	4.57	1.08	3.39	2.67	1.38
2004	2.21	2.77	8.89	5.80	3.75	1.25	4.02	2.62	1.69
2005	1.61	1.88	6.29	3.81	2.61	1.17	3.91	2.36	1.62
2006	0.73	0.85	2.91	1.60	2.21	1.17	3.99	2.19	3.03
2007	0.08	0.09	0.41	0.17	0.00	1.12	5.07	2.14	0.00

Source: HUD; claims paid as of February 29, 2008.

Table 5 can be reproduced using a standard data analysis tool called a regression. Regressions are used to establish the relationship between variables, and can be used to help determine how one variable relates to another. Regressions are widely used by the federal government to analyze programs. For example, the Office of Federal Contract Compliance Programs (OFCCP) is now aggressively using regression as a tool with which they audit contractors.

¹ See Table 5, *Federal Register* 73(116), p. 33953 or click [HERE](#).

Table 5 data is reproduced using a simple linear regression relating only two variables: (1) foreclosure and (2) source of down payment. The problem with Table 5 then, and with almost all of HUD’s tables and analysis, is that it is only valid if the only factor for foreclosures is the source of down payment. This is clearly not the case. There are several important variables that factor into a person’s likelihood of foreclosure. A simple regression clearly supports this. The correlation coefficient, or “R²”, of HUD’s model is only .0067 (R² of 1 being a perfect explanation of the data by the model and 0 being no explanation). So clearly the model HUD chose provides a poor explanation of the relationship between down payment assistance and foreclosures.

When we use the same model that produces Table 5 data for fiscal year 2006,² and then add other variables³ from the data HUD has made public on its website, we observe that the effect of the down payment assistance source on foreclosures declines by over 68%--with the difference between borrower paid down payment assistance and SFDPA falling from 2.18 in HUD’s model to just 1.49. But the R² value is still very low at .0147, which tells us that our model does not provide a comprehensive explanation of the relationship between foreclosures and the use of SFDPA.

When we use a non-linear regression, which is simply another model for regressions, we observe that the effect of the down payment assistance source on foreclosure declines by over 100% from HUD’s model--with the difference between borrower paid down payment assistance and SFDPA falling from 2.18 in HUD’s model to just 1.06. More importantly the R², or how well the model explains the relationship between the variables, increases by over 1400% to .0947, indicating a much better fit of the data. This tells us that not only did HUD fail to account for several important variables, but they likely did not use the best regression model either.

² Adding the coefficients of the regression to the intercept or “constant” yields the results of Table 5—See Column D Below.

Differences between Down payment Source Category and Borrower Category (2006)

Down Payment Source	Column A Coef.	Column B Std. Err.	Column C Coef.*100	Column D Coef.+Const.	Column E t-stat
Employer	0.014814	0.007292	1.48	2.21	2.03
Family	0.001225	0.000506	0.12	0.85	2.42
Gov’t	0.008653	0.000984	0.87	1.6	8.79
Nonprofit	0.02181	0.000567	2.18	2.91	38.44
Borrower (Constant)	0.007299	0.000218	72	0.73	33.54

³ These variables are FICO Category (not FICO score), Loan-To-Value, Race, Original Purchase Price and State.

Because there are several other variables that should be accounted for, accounting for these variables will likely show the effect of SFDPA on foreclosures to be even less. These variables would include: co-borrower's information, collections owed, amount of closing costs, employment, income, debt, monthly payment, bankruptcy, if the borrower is a first-time homebuyer, interest rate, FICO score (not just FICO category), etc.

The important point is that HUD's analysis vastly overstates the effect of SFDPA on foreclosure by pretending that it is the only factor that could affect foreclosure. The Senate should be aware of the flaw in this approach. It is not a well-considered policy to ban SFDPA when all of the statistics supporting such a ban are terribly flawed. It is also a mistake to deny needed assistance to borrowers and the market in general when, even if a proper analysis shows the use of SFDPA does impose additional risk, that risk can so easily be mitigated by imposing a few underwriting standards, such as minimum FICO, Debt/Income ratio, etc.