
1. Questionnaire changes in 2007.

Information from two sections, the housing section (A) and the wealth section (W), were used to construct the 2007 wealth file. The questionnaire part that asks this information has not been changed between the 2005 and 2007 waves.

2. Data processing and imputations.

PSID asks about nine broad wealth categories, including short-term debt:

1. Equity in business (also includes farm) [W11, S803]
2. Transaction accounts (includes savings accounts, money market funds, certificates of deposit, government savings bonds, and treasury bills) [W28, S805]
3. Value of debt aside from mortgage on the main home or vehicle loans [W39, S807]
4. Equity in real estate (second home, land, rental real estate, or money owed on a land contract) [W2, S809]
5. Equity in stock (includes shares of stock in publicly held corporations, mutual funds, and investment trusts) [W16, S811]
6. Equity in vehicle (cars, trucks, a motor home, a trailer, or a boat) [W6, S813]
7. Other assets (includes bond funds, cash value in a life insurance policy, a valuable collection for investment purposes, and rights in a trust or estate) [W34, S815]
8. Equity in Individual Retirement Accounts [W22, S819]
9. Value of home equity (calculated as home value minus remaining mortgage; used in calculation of WEALTH2, but not released separately until March 2011) [S820]

Questions about home equity (primary residence) are asked in section A.
Processing of the data includes three steps: a) imputation of the wealth components (1-8), b) computation of home equity (9), and c) construction of estimates for the total family wealth with and without housing equity. In the processing of the 2007 wealth data we followed the same approach as in the prior waves. In particular, the hot-deck imputation technique was used for imputation of the missing data in the wealth components (1-8). Details on home equity imputation are given in Table 1.

Table 1. Summary of 2007 Home Equity Imputations.

<table>
<thead>
<tr>
<th>Missing data group</th>
<th>Imputation category</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: House value is not missing; mortgage is missing</td>
<td></td>
<td>40</td>
<td>17</td>
<td>94</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>151</td>
</tr>
<tr>
<td>Group 2: House value is missing; mortgage is not missing</td>
<td></td>
<td>25</td>
<td>3</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>Group 3: House value is missing; there is no mortgage</td>
<td></td>
<td>35</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>44</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>Group 4: Both house value and mortgage are missing</td>
<td></td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>105</td>
<td>32</td>
<td>94</td>
<td>18</td>
<td>44</td>
<td>9</td>
<td>31</td>
<td>20</td>
<td>353</td>
</tr>
</tbody>
</table>

There are four rows in Table 1 corresponding to four patterns of missing data.

In Group 1, where mortgage is missing but not housing value, we first attempt to replace the missing mortgage value with a value from the last wave (category A). If this is not possible, a value from the wave prior to the last (2003) is used (category B). When this is also not possible, we use an assumed identity, Mortgage=0.6*HouseValue, to obtain an estimate for the mortgage.

When a housing value is missing but mortgage is not (Group 2) we proceed in the following way. First, using information on the "bracketed" or interval-censored responses, an estimate for the house value is calculated as an average of the lower and upper bracket
values available. Then the missing mortgage is obtained as the maximum value of mortgage value reported in the last wave and the bracket estimate (category A). If the last wave information is not available, we use information from the wave prior to the last wave (category B). Finally, if mortgage value was not reported two waves ago then the missing value is imputed as maximum of house value divided by .4 and the house value estimate based on the bracket information (category D).

In the case where house value is missing and there is no mortgage, i.e. mortgage value is zero (Group 3), we first try to use information on the house value from the last wave or, if needed, an earlier wave. When neither of these is available the missing home value is imputed in the same way as in Group 2, i.e. the missing value is assigned the maximum of the house value reported in a previous wave and the estimate obtained with help of the bracket questions (category A and B). When house value information is not available in the previous waves then the house value is assigned as an average of the upper bracket and the lower bracket values of the house value. Further, if the bracket information is not available then the median of the reported house values among those who have no mortgage is assigned.

Group 4 consists of cases with both the house value and mortgage missing. First, using information from two earlier waves, the missing house value is treated as in the same way as in Group 2 and 3, and the missing mortgage is imputed as in the Group1 (category A and B). If information in two preceding waves is not available but the bracket values for house value are available then house value is estimated as an average of the lower and the upper bracket values and home equity is equal to 40% of this estimate, i.e. home equity is equal to 0.4 * (lower bracket + upper bracket)/2 (category G). If the bracket information for the house value is also missing then home equity is assigned the median value of home equity among all known cases (category H).

Table 2. Summary Statistics of 2007 Imputed Home Equity Values.

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>353</td>
<td>107697.66</td>
<td>107200.05</td>
<td>-30000.00</td>
<td>650000.00</td>
</tr>
</tbody>
</table>


The second release of the 2007 wealth dataset contains 38 variables. Accuracy codes were added for each component of total wealth as well as for the two totals (S802A-S820A), to indentify records which have been imputed/changed. In addition, home equity value is now included as a separate release variable (S820).