February 2021

Simulating Economic Relief Payments
Appendix

James Hawkins

Loading a Custom Input File:

Researchers who are interested in forking the Observablehq code behind our interactive and loading their own input file can do so by constructing their data set to include the following variables.

- *agi*: Includes the Adjusted Gross Income of the tax unit.
- *adults*: Includes the number of adults in the tax unit. Single adult filers should be coded as 1, head of household filers as 1.5, and joint filers as 2.
- *adults_eq*: Includes the number of adults for the purposes of calculating the unit's equivalence scale (the generosity of the payment). Single adult and head of household filers should be coded as 1 and joint filers as 2.
- *children*: Includes the number of dependents in the tax unit who qualify as child dependents.
- *adult_children*: Includes the number of dependents in the tax unit who do not qualify as child dependents.
- *ssn*: Includes a dummy variable for whether the head of the tax unit, and their spouse if joint filing, holds a Social Security Number (SSN). Coded as 1 if they do hold an SSN and 0 if not.
- *decile*: Denotes the decile of the tax unit within the income distribution. Tax units in the first decile should be coded as 1, second decile as 2, etc.
- *weight*: Includes the probability weight for the tax unit. If no "raked_weight" variable is included in the input file, click the "Original Weights (2018 CPS) setting for the Weight Specification parameter – this can be renamed by the user.

Relief Payment Calculations:

Our basic analytical approach involves solving for the budget of a hypothetical income support program based on a set of policy parameters and eligibility thresholds. It is useful to first consider the straightforward example of a universal payment, wherein an income support is provided equally to the entire U.S. population (this is popularly referred to as a "universal basic income"). The universal payment for this program can be solved by dividing the budget of the program by the total size of the population of tax units (under the assumption that the payments are administered by the IRS)

$$P = B/U \tag{1}$$

Where P is the program payment, B is the program budget, and U is the population of tax units.

Next, we introduce a complication to this payment structure by examining a non-universal program, which includes a maximum payment for tax units with no income (the "income guarantee") and "phases out" at a constant rate for each additional dollar of a tax unit's income. For instance, a program with a $1,000 income guarantee and a 5% phaseout rate would phase out completely at $20k. A tax unit with $5,000 of income would receive a payment of $750 -- $250 less than the maximum payment. A payment for a tax unit for such a program is given by the following piecewise function.

$$P = \begin{cases} G - I\tau & \text{if } I < T \\ 0 & \text{otherwise} \end{cases} \tag{2}$$
$$\text{where } \tau > 0$$

Where G is the income guarantee (the maximum payment); I is the income of the tax unit; T is an exogenously defined eligibility threshold, at or above which tax units are no longer eligible and receive zero payment; and tau is the positively-defined phaseout rate. Tax units with incomes below the eligibility threshold are in what we refer to as the “phaseout region.” For the purpose of reducing the number of unknown variables in the equation, it is useful to redefine G as the product of T and tau.

$$\begin{gathered} G = T\tau \\ P = \begin{cases} (T - I)\tau & \text{if } I < T \\ 0 & \text{otherwise} \end{cases} \\ \text{where } \tau > 0 \end{gathered} \tag{3}$$

The intuition behind this is given by the fact that the phaseout rate is the rise (G) over the run (T) of the income support program. Thus, G is equal to the product of T and tau.

We can aggregate this expression to the population level to facilitate policy analysis.

$$\begin{gathered} \sum_{i=1}^{n_1} P_i = \sum_{i=1}^{n_1} (T - I_i)\tau \\ \sum_{i=1}^{n_2} P_i = n_2 0 \end{gathered} \tag{4}$$

Where each tax unit is indexed with subscript i, the number of tax units with incomes below the eligibility threshold -- the phaseout region -- is represented by n_1, and the number of tax units with incomes at or above the eligibility threshold is represented by n_2.

$$B \equiv \sum_{i=1}^{n_1} P_i + \sum_{i=1}^{n_2} P_i \tag{5}$$

By definition, the budget of the program, denoted by B, is equivalent to the sum of all payments in the program. As we have defined it, tax units above the eligibility threshold receive zero dollars in payment. Thus, payments for these units sum to zero (i.e., they receive no payments) and this term can be dropped from equation 5, leaving the budget equal to the sum of the payments in the phaseout region.

$$B = \sum_{i=1}^{n_1} (T - I_i)\tau \tag{6}$$

We can rearrange equation 6 after factoring out tau -- since it is a constant -- so that the phaseout rate is a function of the budget and the eligibility threshold. This provides a solution for this income support program given exogenous values for the budget and the eligibility threshold.

$$\tau = \frac{B}{\sum_{i=1}^{n_1} (T - I_i)} \tag{7}$$

We can extend this framework to model relief payments of a similar structure as those enacted via the CARES Act. These payments introduce a “kink point,” with tax units with incomes at or below that point receiving an equivalent maximum payment. This is depicted in Figure A1 as the “payment guarantee” region. Any tax unit with income above the kink point experiences a constant phaseout of their payment for the portion of their income greater than the kink point. This is similar to the phaseout in Equation 3; however, the phaseout rate is not applied to a tax unit’s entire income -- it only applies to that portion above the kink point.

$$P = \begin{cases} G & \text{if } I \leq K \\ G - (I - K)\tau & \text{if } I > K \text{ and } I < T \\ 0 & \text{otherwise} \end{cases} \tag{8}$$
$$\text{where } \tau > 0$$

Where K is defined as the kink point. As with the previous iteration of the relief program model, it is useful to convert G into a function of tau to reduce the number of unknowns. As with Equation 3, G is the rise over the run; however, since we now have an entire region where the payment is flat, we need to subtract the kink point from the eligibility threshold to calculate the proper length of the run, which excludes the distance between zero and the kink point (see Figure A1).

$$G = (T - K)\tau$$
$$P = \begin{cases} (T - K)\tau & \text{if } I \leq K \\ (T - K)\tau - (I - K)\tau & \text{if } I > K \text{ and } I < T \\ 0 & \text{otherwise} \end{cases} \tag{9}$$
$$\text{where } \tau > 0$$

As with the previous program model, we can aggregate this expression to the population level to facilitate policy analysis.

$$
\begin{aligned}
\sum_{i=1}^{n_1} P_i &= n_1(T-K)\tau \\
\sum_{i=1}^{n_2} P_i &= n_2(T-K)\tau - \tau\sum_{i=1}^{n_2}(I_i - K) \\
\sum_{i=1}^{n_3} P_i &= n_3 0
\end{aligned}
\tag{10}
$$

Where each tax unit is again indexed with subscript i, the number of tax units with incomes at or below the kink point -- the payment guarantee region -- is represented by n_1; the number of units with incomes above the kink point and below the eligibility threshold, the new phaseout region, is represented by n_2; and the number of tax units with incomes at or above the eligibility threshold is represented by n_3. Since G is a constant (represented by T minus K, multiplied by tau in Equation 10), we can multiply it by the number of tax units in each of the eligible segments of the program, n_1 and n_2, in Equation 10.

$$
B \equiv \sum_{i=1}^{n_1} P_i + \sum_{i=1}^{n_2} P_i + \sum_{i=1}^{n_3} P_i \tag{11}
$$

By definition, the budget is equal to the sum of all payments. Since the payments for tax units above the eligibility threshold are equal to zero, we can drop this from Equation 11 to yield our complete budget constraint.

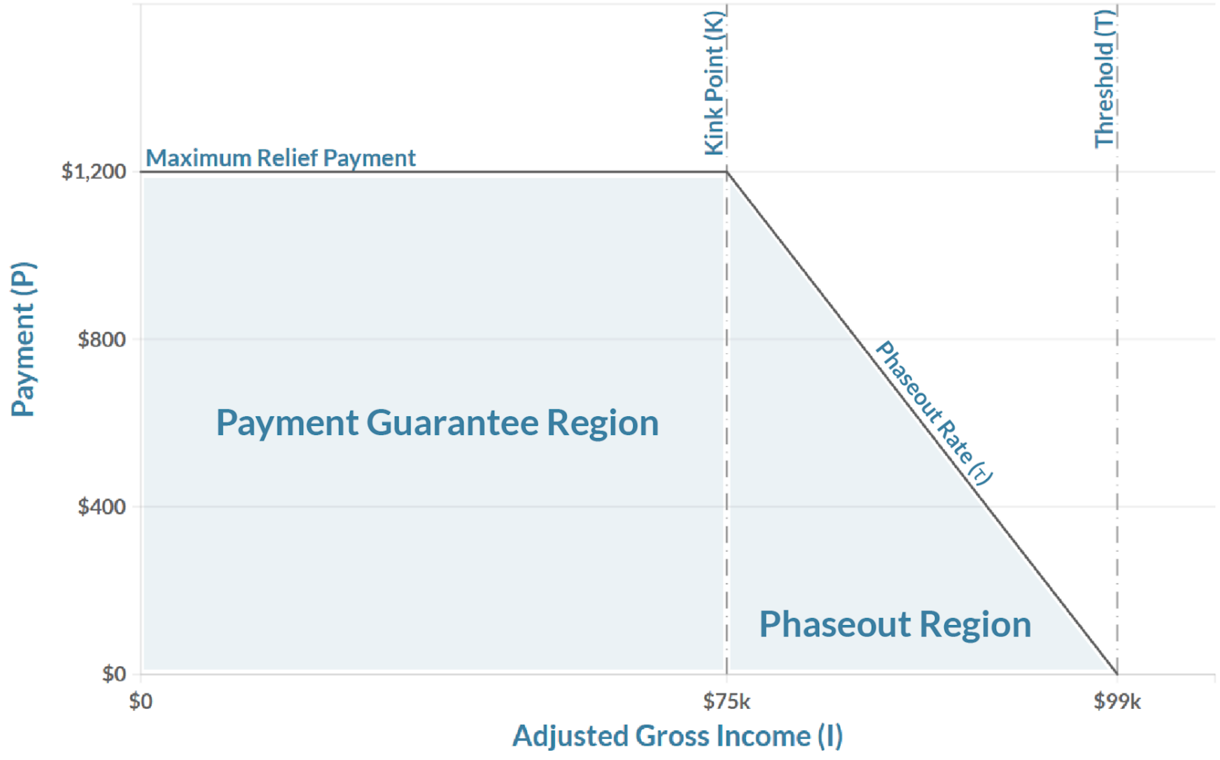
$$
B = n_1(T-K)\tau + n_2(T-K)\tau - \tau\sum_{i=1}^{n_2}(I_i - K) \tag{12}
$$

Finally, we can factor out the tau's from Equation 12 and rearrange to yield a new formula for the phaseout rate as a function of the budget, the eligibility threshold, and the kink point.

$$
\tau = \frac{B}{n_1(T-K) + n_2(T-K) - \sum_{i=1}^{n_2}(I_i - K)} \tag{13}
$$

We can then use the phaseout rate calculated in equation 13, given values for the budget, eligibility threshold, and kink point, and plug it back into equation 9, to yield the relief payment for each tax unit. This formula can be extended to account for tax units of different sizes (equivalence scales), differential phaseout rates for different filing categories, and categorical eligibility conditions, as we do in the code of our simulation.

Figure A1: Illustrative Payment Schedule



Notes: Graph not drawn to scale.

Raked Weights:

We reweight our sample of tax units that we identify as eligible for CARES Act relief payments based on a set of administrative targets from data published by the IRS. Specifically, we use the raking algorithm implemented by the Stata ipfraking package to make adjustments to the original CPS weights to match administrative aggregates published by the IRS for the total number of relief payments in eight Adjusted Gross Income brackets, total number of relief payments in each state, total number of relief payments in each state for tax units with no qualifying dependents, and total number of relief payments in each state for tax units with qualifying dependents. We remove payments to eligible residents not residing at a U.S. state address from the state-level targets (according to the IRS, this includes residents of U.S. territories, U.S. citizens living abroad, and U.S. armed forces overseas. The IRS does not report the number of payments going to eligible residents at non-state addresses at the AGI level; therefore, we calculate the share of payments going to non-state addresses based on the state-level data and adjust the AGI bracket-based targets downward uniformly based on this state-calculated figure.

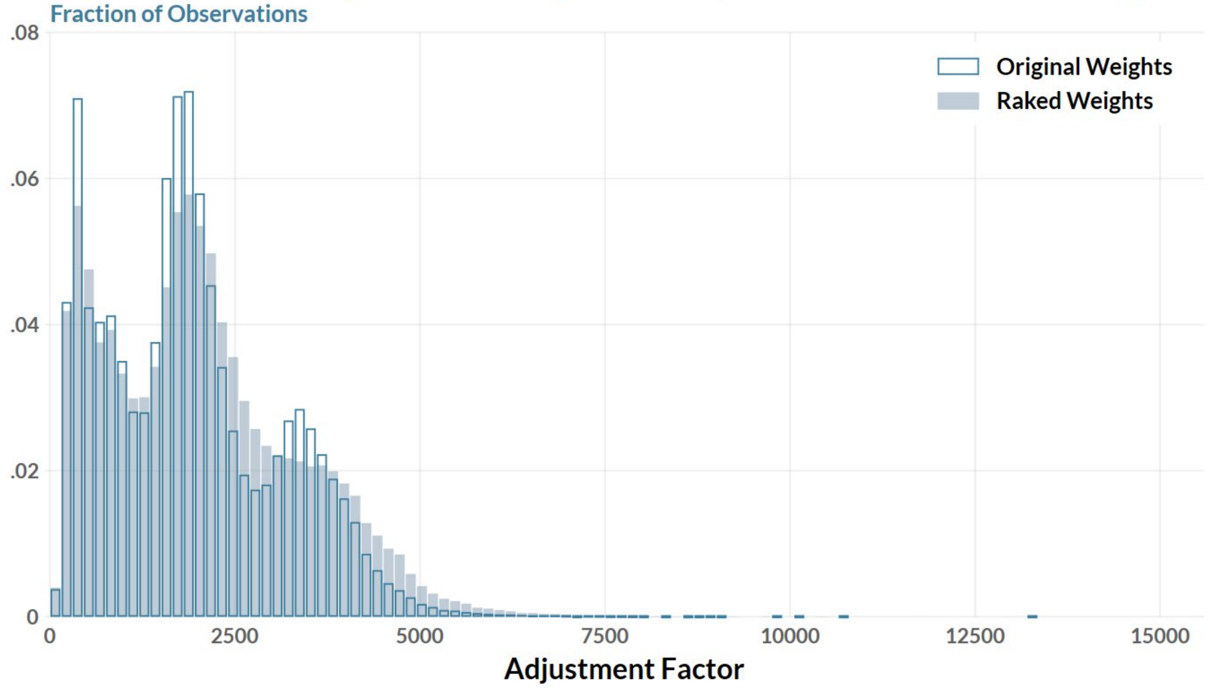
The raking algorithm generally upweights underreporting of relief payments in lower income brackets and down-weights overreporting of relief payments in the highest income brackets (see “Benchmarks” in the Appendix). The raked weight and the original weight are equivalent for tax units who we simulated as ineligible for the CARES Act relief payments since we do not have administrative targets for these units.

In Figure A2 and A3, we show how the distribution of the raked weights differ from the original weights and the distribution of adjustment factors (the relative difference between the raked weights and the original weights). The peak in the distribution of adjustment factors to the original weight occurs between 1.14 and 1.18, with 90% of the observations occurring between an adjustment factor of .84 and 1.298.

Adjustments to Budget:

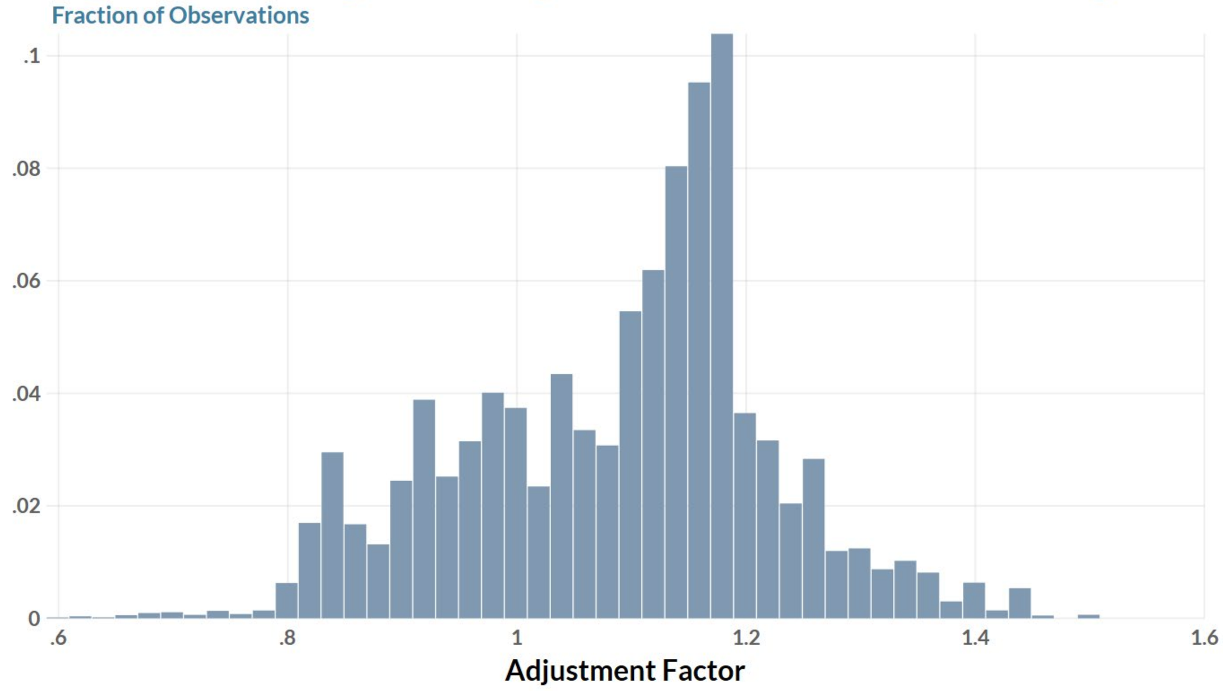
We make adjustments to the total program budget to bring our simulated estimates in closer alignment with real-world outcomes for our sample of residents of U.S. states. Specifically, our sample does not include data on residents eligible for the CARES Act residing at non-U.S. state addresses. Based on data published by the IRS, relief payments were sent to 1.18 million tax units for an aggregate of $1.8 billion relief payments paid to eligible tax units residing at a non-state address. This $1.8 billion accounts for approximately .6635% of the $271.4 billion aggregate payments for the CARES Act. We adjust the available share of the total user-inputted budgets downward to reflect the share of the budget for payments made to units at non-state addresses. For example, a user-inputted budget of $300 billion would be adjusted downward by $1.9 billion so that $298.1 billion was available for relief payments to state-based tax units.

Figure A2: Histogram of Original Weights Versus Raked Weights



Source: Author's analysis March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org), SSN status imputed by the author, and raked weights computed based on IRS targets.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket. Raked weights are calculated using the ipfraking Stata command based on a series of target program totals published by the IRS for different AGI brackets and at the state level.

Figure A3: Histogram of Adjustment Factor for Raked Weights



Source: Author's analysis March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org), SSN status imputed by the author, and raked weights computed based on IRS targets.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket. Raked weights are calculated using the ipfraking Stata command based on a series of target program totals published by the IRS for different AGI brackets and at the state level.

Benchmarks:

We conduct benchmarking of the accuracy of our simulation model using published data of CARES Act relief payments from the IRS. Generally, we find that our simulation of the CARES Act with the original CPS weights underreports the number of relief payments for tax units with lower incomes and overreports the number of relief payments for tax units with the highest incomes. After adjusting for the share of payments going to eligible tax units residing at foreign addresses, we find that our simulation of aggregate number of payments with the original CPS weights is about 91.8% of the IRS figure. Our simulation of aggregate expenditures with the original weights is approximately 97.7% of the IRS figures. Estimates of the aggregate number of relief payments to eligible residents of U.S. states using the raked weights are 100% of the same measure from the IRS. Our simulation of aggregate expenditures using the raked weights is 106.2% of IRS figures, after excluding payments to non-state tax units.

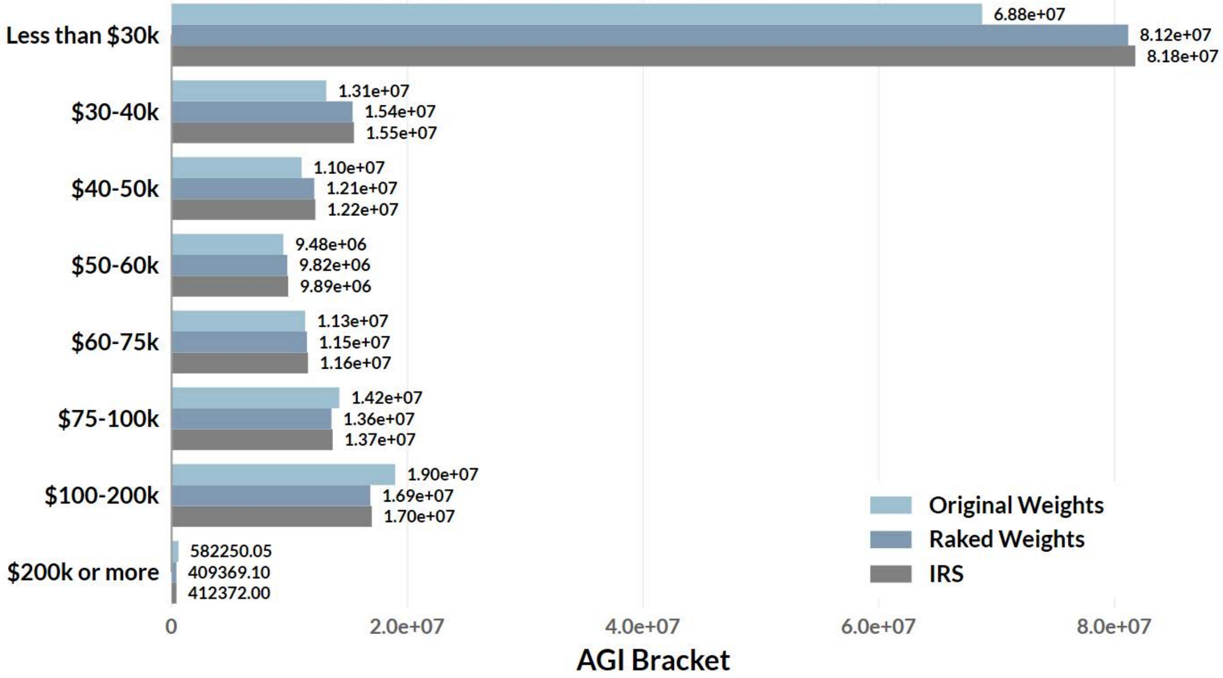
The typical payment in our simulation is greater than those reported by the IRS for the CARES Act. Therefore, once we implement our raked weights to match the IRS targets for the aggregate number of payments, our simulation will tend to overestimate total expenditures relative to the IRS. Based on this benchmarking, we generally expect that any simulation using our calculator will overestimate the program budget by about 6%. Put in different terms, our simulation will underreport the dollar amount of relief payments possible with a given budget, relative to what IRS reports for the CARES Act. This bias may differ depending on the eligibility settings chosen in the interactive calculator. For instance, if the components of Adjusted Gross

Income, which is a key input for our calculator, are misreported for low- or upper-incomes in the CPS, then simulations that are highly targeted based on income or more universal may not approximate real-world outcomes.

In Figures A4 through A7 we show benchmarks of our simulation of the CARES Act for different Adjusted Gross Income brackets. We group all tax units with AGI below $30k into a single bracket and include non-filers and units with negative AGI in this category to avoid the need for imputations of who are filers/non-filers in the CPS data. Unlike the aggregate benchmarks we discuss above, we do not adjust the IRS data used in Figures A4 through A7 for the number of payments going to eligible residents at foreign addresses. Thus, the data shown in these figures uses the raw data provided by the IRS with no adjustments. The discrepancies shown in Figure A4 between the number of payments using the raked weights and the IRS data is due to this lack of adjustment.

Figure A4: Comparing Aggregate Count of Payments

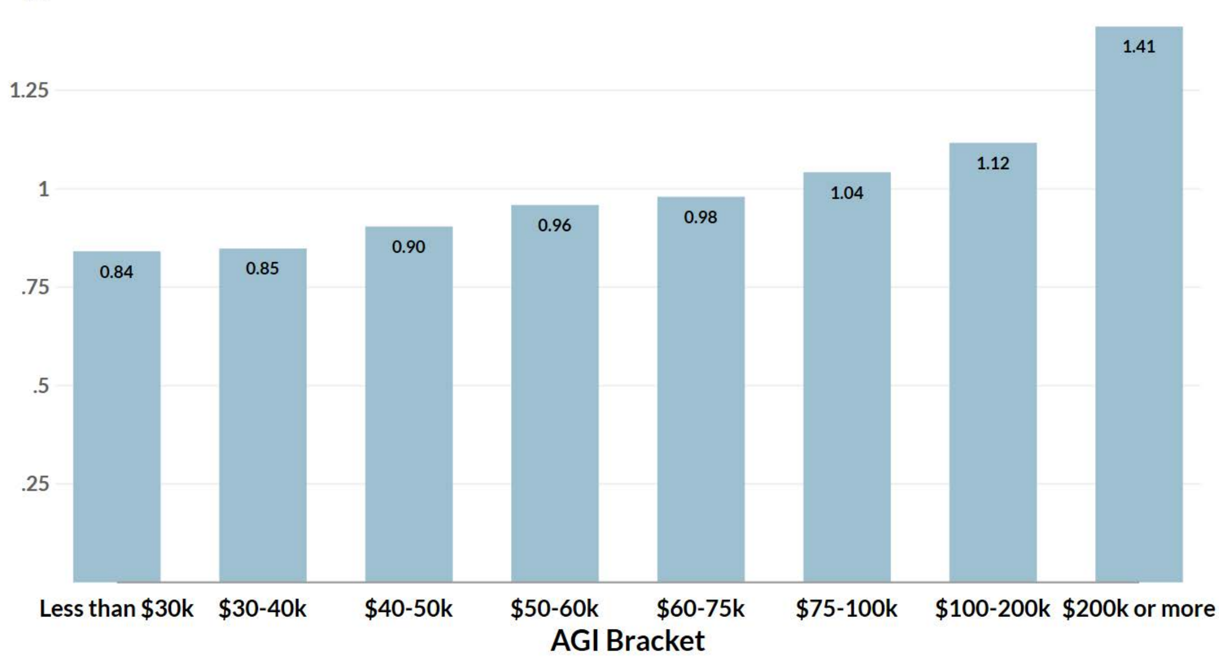
Simulation Aggregates of CARES Act Versus IRS Reported Aggregates



Source: Author's analysis of aggregated data from the IRS and March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org) and SSN status imputed by the author.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket.

Figure A5: Comparing Aggregate Count of Payments

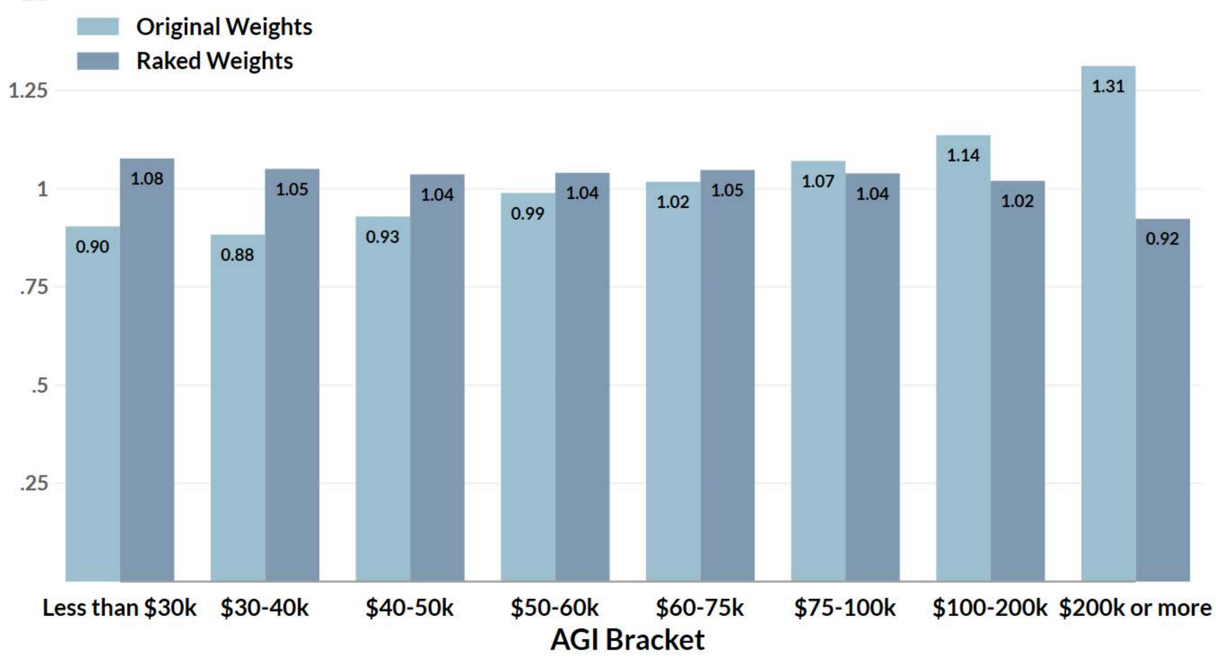
Simulation Aggregates of CARES Act as a Fraction of IRS Reported Aggregates (Original Weights)



Source: Author's analysis of aggregated data from the IRS and March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org) and SSN status imputed by the author.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket. Raked weights are calculated using the ipfraking Stata command based on a series of target program totals published by the IRS for different AGI brackets and at the state level.

Figure A6: Comparing Aggregate Dollar Amount of Payments

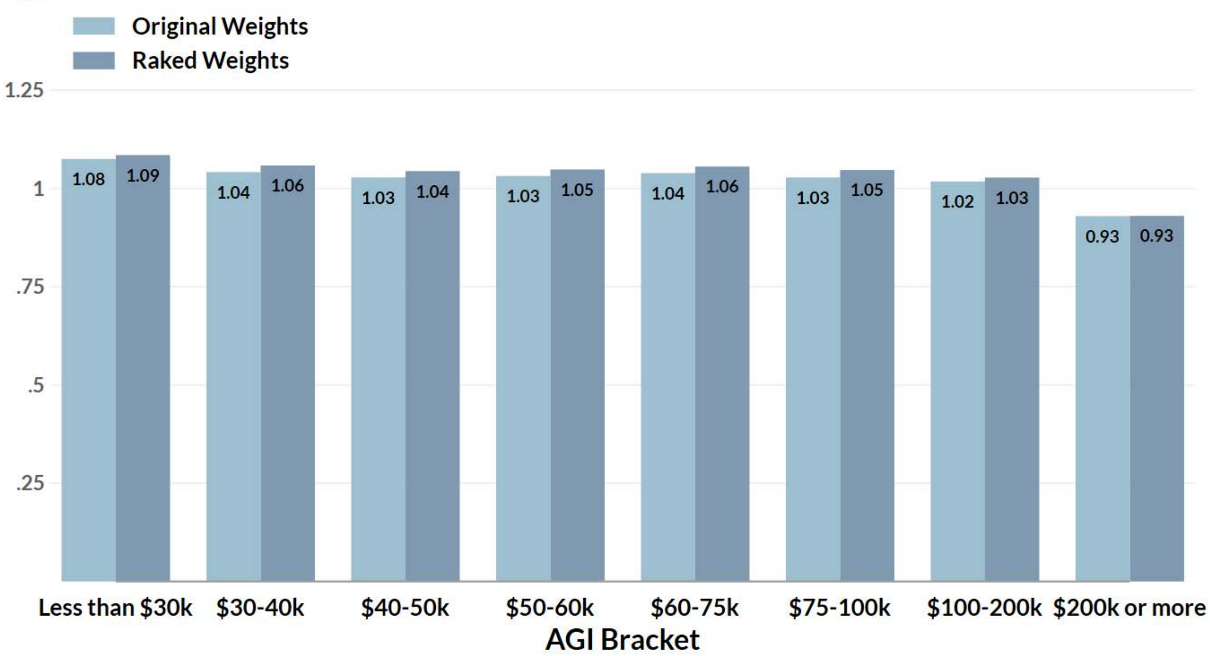
Simulation Aggregates of CARES Act as a Fraction of IRS Reported Aggregates



Source: Author's analysis of aggregated data from the IRS and March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org) and SSN status imputed by the author.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket. Raked weights are calculated using the ipfraking Stata command based on a series of target program totals published by the IRS for different AGI brackets and at the state level.

Figure A7: Comparing Average Payments

Simulation Averages of CARES Act as a Fraction of IRS Reported Averages



Source: Author's analysis of aggregated data from the IRS and March 2018 CPS ASEC data augmented with tax unit identification produced by the TRIM3 microsimulation model (trim3.urban.org) and SSN status imputed by the author.
Notes: IRS reported figures for non-filers and tax units with negative AGI are included in the <$30k bracket. Raked weights are calculated using the ipfraking Stata command based on a series of target program totals published by the IRS for different AGI brackets and at the state level.