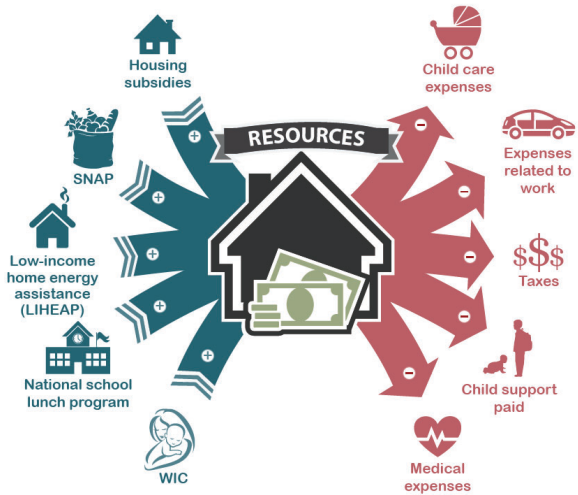
# An Info-Metrics Approach to Estimating the Supplemental Poverty Rate of Public Use Microdata Areas

Danielle Wilson

American University

March 30, 2021

# Supplemental Poverty Measure



# Supplemental Poverty Measure (SPM)

- An individual's poverty status is defined as:

$$P_i = \begin{cases} 1, & \boldsymbol{R_i} < \boldsymbol{T_i} \\ 0, & \text{o.t.} \end{cases} \tag{1}$$

where $\boldsymbol{T_i}$ is a Bureau of Labor Statistic's defined threshold and

$$\begin{aligned} \boldsymbol{R}_i &= \sum_{j=1}^{13} r_{ij} \\ &= \mathit{FamilyCashIncome}_i - \mathit{ChildCareExpenses}_i - \mathit{MedicalExpenses}_i \\ &\quad - \mathit{ChildSupport}_i + \mathit{EnergySubsidy}_i + \mathit{SNAP}_i - \mathit{FederalTaxes}_i \\ &\quad - \mathit{StateTaxes}_i - \mathit{FICA}_i - \mathit{WorkExpenses}_i + \mathit{HousingSubsidy}_i \\ &\quad + \mathit{WIC}_i + \mathit{SchoolLunch}_i \end{aligned} \tag{2}$$

- Individuals in the same SPM family have the same poverty status.

# Supplemental Poverty Measure (SPM)

- Uses the Current Population Survey's Annual Social and Economic Supplement (CPS-ASEC).
- Unlike regular months of the CPS, the supplement (given in March) asks for detailed income data.
- Produces estimates at the national and state level (using three-year averages).

The Supplemental Poverty Measure: 2019

Current Population Reports

By Liana Fox
P60-272
September 2020

INTRODUCTION

Since the publication of the first official U.S. poverty estimates, researchers and policymakers have continued to discuss the best approach to measure income and poverty in the United States. Beginning in 2011, the U.S. Census Bureau began publishing the Supplemental Poverty Measure (SPM), which extends the official poverty measure by taking account of many of the government programs designed to assist low-income families and individuals that are not included in the official poverty measure. The SPM is produced with the support of the Bureau of Labor Statistics (BLS), and this is the tenth in the series. This report presents estimates of the prevalence of poverty in the United States using the official measure and the SPM based on information collected in 2020 and earlier Current Population Survey Annual Social and Economic Supplements (CPS ASEC). The data collection period for the 2020 CPS ASEC coincided with the COVID-19 pandemic, the associated public health response, and the end of the economic expansion. For details on the impact of COVID-19 on CPS ASEC data collection, see the text box "The Impact of the Coronavirus (COVID-19) Pandemic on the CPS ASEC."

Figure 1.
Supplemental Poverty Measure Rates for Total Population and by Age Group: 2018 and 2019
(in percent)




Note: For information on confidentiality protection, sampling error, nonsampling error, and definitions, see <www.census.gov/programs-surveys/cps/technical-documentation.html>.
Source: U.S. Census Bureau, Current Population Survey, 2018 and 2020 Annual Social and Economic Supplements (CPS ASEC).

HIGHLIGHTS

- In 2019, the overall SPM rate was 11.7 percent. This was 1.0 percentage point lower than the 2018 SPM rate of 12.8 (Figure 1).[1]

[1] Calculated differences here and throughout this report may differ due to rounding.

United States® Census Bureau

U.S. Department of Commerce
U.S. CENSUS BUREAU
census.gov

# Supplemental Poverty Measure (SPM) with the American Community Survey (ACS)

- In 2020, Fox, Glassman and Pacas published a working paper detailing how the SPM can be used in the ACS.

+ "Census Bureau recommends the use of the American Community Survey (ACS) for poverty estimates for sub-national geographic units" because of its **larger sample size**.

– However, the "ACS is **missing a number of key data elements** required to produce SPM estimates".

# Comparing the Supplemental Poverty Measure Between Surveys

## CPS-ASEC (Official Measure)

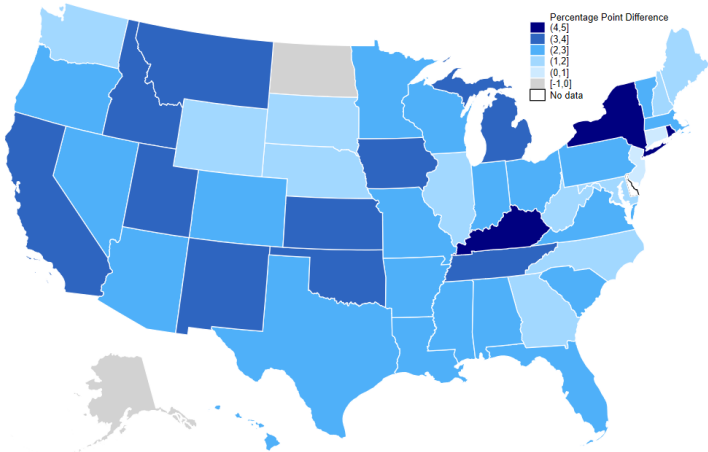
+ Detailed income data.
− Small sample size - **only** state estimates.

## ACS

+ Large sample size - **state & PUMA** estimates.
− Less detailed data (requires imputation of more components).

# Comparing the Supplemental Poverty Measure Between Surveys

Difference Between the CPS-ASEC and ACS SPM Rates by State: 2016 - 2018



# Proposed Method

**Given that the SPM ACS rates require more imputation, why "throw away" or disregard more accurate information from the CPS-ASEC?**

- Inspired by Papalia & Fernandez-Vasquez's (2020) information-theoretic solution using count data, will similarly **constrain disaggregate estimates with** more reliable, out of sample, **aggregate information**.
  - *In sample data* will come from the ACS, that can produce *PUMA estimates*.
  - *Out of sample data* will be the *state estimates* from the CPS-ASEC.

# Comparing the Supplemental Poverty Measure Between Surveys

## CPS-ASEC (Official Measure)

Directly Observe:

* Family Cash Income
* Child Care Expenses
* Medical Expenses
* Energy Subsidy
* SNAP

Require Imputation:

- Federal Taxes
- State Taxes
- FICA
- Workers Expenses
- Housing Subsidy
- WIC
- School Lunch

## ACS

Directly Observe:

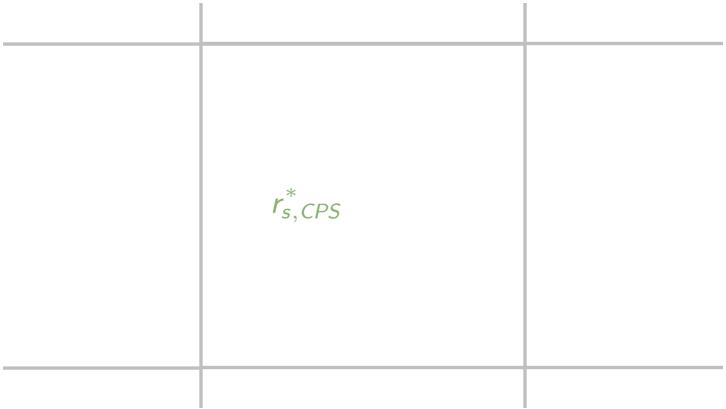
* Family Cash Income

Require Imputation/Out-of-Sample Data:

- Child Care Expenses
- Medical Expenses
- Federal Taxes
- State Taxes
- FICA
- Workers Expenses
- Energy Subsidy
- SNAP
- Housing Subsidy
- WIC
- School Lunch

# Applying the Proposed Method

State Level Supplemental Poverty Rate in the CPS-ASEC

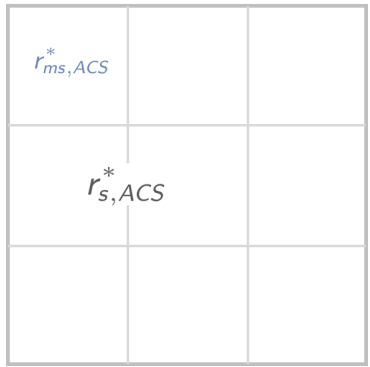


# Applying the Proposed Method

## State Level Supplemental Poverty Rate in the ACS

# Applying the Proposed Method

## PUMA and State Supplemental Poverty Rates in the ACS



# Applying the Proposed Method

## PUMA and State Supplemental Poverty Rate in the ACS

| | | |
|---|---|---|
| $r^*_{ms,ACS}$ | 2 | 3 |
| 4 $r^*_{s,ACS}$ | 5 | 6 |
| 7 | 8 | 9 |

$$r^*_{s,ACS} = \frac{\sum_{m=1}^{M} r^*_{ms,ACS} \times N_{ms,ACS}}{N_{s,ACS}}$$

# Applying the Proposed Method

## PUMA and State Supplemental Poverty Rate in the ACS

| | | |
|---|---|---|
| $r^*_{ms,ACS}$ | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

$$r^*_{ms,ACS} = r_{sm,ACS} + \mu_{sm,ACS}$$

# Applying the Proposed Method

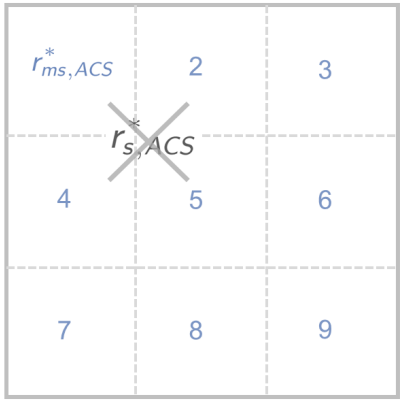
## PUMA and State Supplemental Poverty Rate in the ACS


$$\underbrace{r^*_{ms,ACS}}_{} = r_{sm,ACS} + \mu_{sm,ACS}$$

$$r^*_{s,ACS} = \frac{\sum_{m=1}^{M} \overbrace{(r_{sm,ACS} + \mu_{sm,ACS})}^{r^*_{ms,ACS}} \times N_{ms,ACS}}{N_{s,ACS}}$$

# Applying the Proposed Method

## PUMA and State Supplemental Poverty Rate in the ACS




$$\cancel{r^*_{s,ACS}} = \frac{\sum_{m=1}^{M} \overbrace{(r_{sm,ACS} + \mu_{sm,ACS})}^{r^*_{ms,ACS}} \times N_{ms,ACS}}{N_{s,ACS}}$$

# Applying the Proposed Method

## Relating ACS PUMA Poerty Rates to its CPS State-level Counterpart


$$r^*_{s,CPS} = \frac{\sum_{m=1}^{M} \overbrace{(r_{sm,ACS} + \mu_{sm,ACS})}^{r^*_{ms,ACS}} \times N_{ms,ACS}}{N_{s,ACS}}$$

# Setting Up the Model

*More accurate piece of information constraining the less accurate piece of information:*

$$r^*_{s,CPS} = \frac{\sum_{m=1}^{M} \left( r_{sm,ACS} + \overbrace{\mu_{sm,ACS}}^{Noise} \right) \times N_{ms,ACS}}{N_{s,ACS}} \tag{3}$$

- Goal of the info-metrics model will be to estimate the noise in the constraint, $\mu_{sm,ACS}$.
- $\hat{\mu}_{sm,ACS}$ will provide an estimate of how additional imputation in the ACS may be biasing the estimate **relative to the CPS-ASEC**.
- Will estimate the noise according to the maximum entropy criteria.
- Why maximum entropy criteria? Want to most conservative estimate relative to the information we have - i.e., equation (11).

# Setting Up the Model

- In order to estimate the noise, need to redefine it in discrete terms:

$$\underbrace{\mu_{sm,ACS}}_{\text{Noise}} = \underbrace{\sum_{e=1}^{3} \overbrace{w_{esm,ACS}}^{Weight} \overbrace{v_{esm,ACS}}^{\text{Support element}}}_{\text{Weighted average of support space}}$$

I define the support space using the S.E.s from the (weighted) survey estimates:

$$\boldsymbol{v_{sm,ACS}} = (v_{1sm}, v_{2sm}, v_{3sm}) = (-x * SE, 0, x * SE)$$

▸ Support Space Details

# The Model

The maximum entropy problem for each state $s$, across $M$ respective PUMAs is:

$$\max_{\boldsymbol{w}} H(W) = -\sum_{m=1}^{M}\sum_{p=1}^{3} w_{psm} log(w_{psm})$$

subject to

$$r^*_{s,CPS} = \frac{\sum_{m=1}^{M}\left[r^*_{sm,ACS} - \sum_{e=1}^{3} w_{psm} v_{psm}\right] N^*_{sm,ACS}}{N^*_{s,ACS}} \tag{3}$$

$$\sum_{p=1}^{3} w_{psm,ACS} = 1 \ , \ w_{pms_ACS} \geq 0 \tag{4}$$

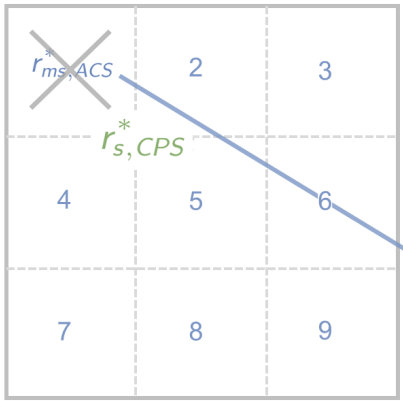
$$1 \geq \sum_{m=1}^{M}\left[r^*_{sm,ACS} - \sum_{p=1}^{3} w_{psm} v_{psm}\right] \geq 0 \tag{5}$$

# Output from the Model


The max. entropy model produces $\hat{w}_{sm,ACS}$ for every PUMA

$$\underbrace{\hat{\mu}_{sm,ACS}}_{\text{Estimated relative noise}} = \sum_{e=1}^{3} \hat{w}_{esm,ACS}\, v_{esm,ACS}$$

# Output from the Model




The max. entropy model produces $\hat{w}_{sm,ACS}$ for every PUMA:

$$\underbrace{\hat{\mu}_{sm,ACS}}_{\text{Estimated relative noise}} = \sum_{e=1}^{3} \hat{w}_{esm,ACS}\, v_{esm,ACS}$$

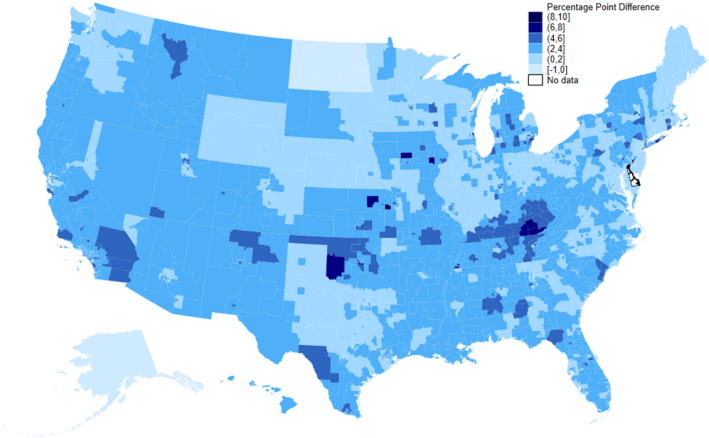
Convert into percentage point form then:

$$\hat{r}_{sm,ACS} = \underbrace{r^*_{ms,ACS}}_{\text{Observed}} - \underbrace{\hat{\mu}_{sm,ACS}}_{\text{Estimated relative noise}}$$

# Results

Difference Between the Max. Entropy and Original ACS SPM Rates by PUMA: 2016 - 2018

$$\hat{\mu}_{sm,ACS} = r^*_{sm,ACS} - \hat{r}_{sm,ACS}$$

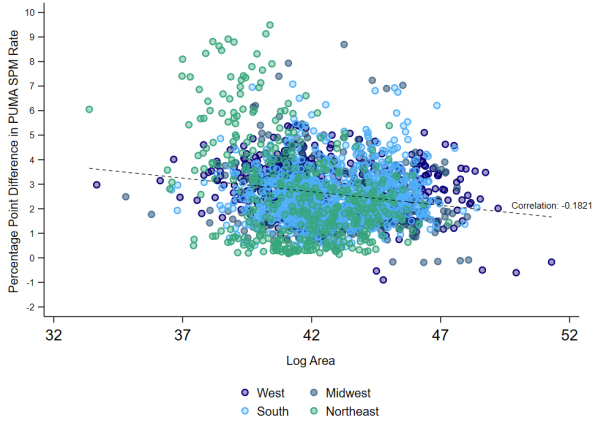


# Results

- Any patterns in estimated error by region or among urban PUMAs?
- No strong patterns between size of PUMA and max. entropy estimate (all PUMAS contain 200k - 300k people).

Difference Between the Max. Entropy and Original ACS SPM Rates by PUMA: 2016 - 2018

$$\hat{\mu}_{sm,ACS} = r^*_{sm,ACS} - \hat{r}_{sm,ACS}$$



# Results

- Area and population have little meaningful impact on max. entropy estimate.
- State controls eliminate substantial amount of prediction error.
- Heteroskedastic errors in model - suggests OVB and that state controls are possibly capturing other state specific characteristics *that may be consequential to the supplemental poverty rate*.

$$\hat{\mu}_{sm,ACS} = \alpha + \gamma_{1,sm} ln(area_{sm}) + \gamma_{2,sm} ln(pop_{sm}) + \gamma_{3,sm} ln(r^*_{sm,ACS}) + \beta \boldsymbol{R_s} + \epsilon_{sm}$$

Table: Magnitude of PUMA-level Refinement: 2016 - 2018

| Regional Controls: | Census Regions (1) | State (2) |
|---|---|---|
| Log Geographic area | 0.00757 | -0.0513*** |
| | (0.0105) | (0.00742) |
| Log Poulation | 0.552*** | 0.543*** |
| | (0.0981) | (0.0077) |
| SPM Rate (Perct. Points) | 0.0882*** | 0.0831*** |
| | (0.00859) | (0.00690) |
| West | -0.467*** | |
| | (0.0830) | |
| Midwest | -0.225** | |
| | (0.0630) | |
| South | -0.630*** | |
| | (0.0786) | |
| Cons. | -4.734*** | -2.227** |
| | (0.956) | (0.783) |
| N | 2,345 | 2,345 |
| adj. R-sq | 0.409 | 0.755 |
| Root MSE | 1.0396 | .66921 |

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Robust standard errors in parentheses. Dependent variable is in percentage points. PUMAs in the state of DE not included.

# Conclusion

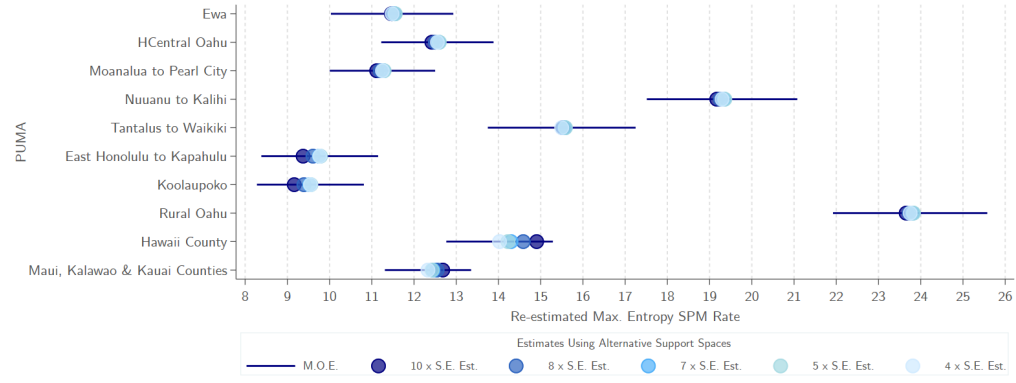
- Info-metrics approach allow for the reconciliation of the same type of information from different sources.
- Each source has pros and cons - but can use the pros from each in a single model.
- Model produces *refined* supplemental poverty rates at the PUMA level.
  - refined rates are limited - subject to CPS-ASEC data, which is better but too naturally imperfect.
- Should be interpreted as conservative, alternative supplemental poverty rates.
- A complement to original ACS SPM rates that facilities poverty and policy analysis at the micro-data level.

Thank you

Questions?

# Appendix Slides

# Max. Entropy PUMA Supplemental Poverty Estimates in Hawaii Using Different Support Spaces




Note: Margin of error (M.O.E.) is for the 99 percent confidence interval.

▸ Return to Slide

# Table: Robustness Checks on Support Space

| Support Space 4SE | | | | | |
|---|---|---|---|---|---|
| 90% C.I. | | 95% C.I. | | 99% C.I. | |
| State | Num. PUMAs | State | Num. PUMAs | State | Num. PUMAs |
| PENNSYLVANIA | 18 | PENNSYLVANIA | 5 | CALIFORNIA | 1 |
| CALIFORNIA | 7 | CALIFORNIA | 3 | MISSOURI | 1 |
| MISSOURI | 3 | MISSOURI | 2 | PENNSYLVANIA | 1 |
| WISCONSIN | 2 | WISCONSIN | 2 | WISCONSIN | 1 |
| ARIZONA | 1 | ARIZONA | 1 | | |
| ARKANSAS | 1 | INDIANA | 1 | | |
| HAWAII | 1 | | | | |
| INDIANA | 1 | | | | |
| OREGON | 1 | | | | |
| VIRGINIA | 1 | | | | |

| Support Space 5SE | | | | | |
|---|---|---|---|---|---|
| 90% C.I. | | 95% C.I. | | 99% C.I. | |
| State | Num. PUMAs | State | Num. PUMAs | State | Num. PUMAs |
| IOWA | 7 | IOWA | 6 | IOWA | 2 |
| TENNESSEE | 3 | CALIFORNIA | 1 | CALIFORNIA | 1 |
| MISSOURI | 2 | MONTANA | 1 | OKLAHOMA | 1 |
| OKLAHOMA | 2 | OKLAHOMA | 1 | | |
| CALIFORNIA | 1 | TENNESSEE | 1 | | |
| MONTANA | 1 | UTAH | 1 | | |
| PENNSYLVANIA | 1 | | | | |
| UTAH | 1 | | | | |
| WISCONSIN | 1 | | | | |

| Support Space 7SE | | | | | |
|---|---|---|---|---|---|
| 90% C.I. | | 95% C.I. | | 99% C.I. | |
| State | Num. PUMAs | State | Num. PUMAs | State | Num. PUMAs |
| KANSAS | 1 | OKLAHOMA | 1 | | |
| MICHIGAN | 1 | | | | |
| OKLAHOMA | 1 | | | | |

▸ Return to Slide