Evaluation of Population Projection Errors

Mathew E. Hauer

February 27, 2018

Overview

The cohort-component method is the most accepted methodology to produce population projections. The method makes use of all three population component processes (fertility, mortality, and migration) and applies them across varying population cohorts to arrive at a future population. Equation 1 outlines the basic structure of a cohort-component model.

$$P_{t+1} = P_t + B_t - D_t + M_{t,in} - M_{t,out} \tag{1}$$

Where P_t is the population at time t, B_t is the births at time t, D_t is the deaths at time t, and $M_{t,in/out}$ refers to in- or out-migration at time t.

Cohort-component requires data on each component process disaggregated by age, sex, and race. Certain elements of these data can be difficult to obtain for national coverage. Birth and death data are typically obtained through the National Center of Health Statistics (NCHS) vital events registration databases. These data, however, are only available for counties with populations greater than 100k and are suppressed in populations with fewer than 1k (I think) members rendering a universal county-level population projection difficult, if not impossible, to complete using publicly available datasets.

An alternative to cohort-component is the Hamilton-Perry method, which uses cohort-change ratios (CCRs) in place of components to project populations. The basic CCR equation is found in equation 2.

$$\begin{aligned} CCR_t &= \frac{{}_nP_{x,t}}{{}_nP_{x-y,t-1}} \\ {}_nP_{x+t} &= CCR_t \cdot {}_nP_{x-y,t} \end{aligned} \tag{2}$$

Where ${}_nP_{x,t}$ is the population aged x to $x+n$ in time t and ${}_nP_{x-y,t}$ is the population aged x to $x+n-y$ in time t where y refers to the time difference between time periods. These CCRs are calculated for each age group a, for each sex group s, for each race group r, in each time period t, in county c. Thus to find the population of ten to fourteen year olds (${}_5P_{10}$) in five years ($t+1$), we multiply the ratio of the population aged 10-14 in time t (${}_5P_{10,t}$) to the population aged 5-9 five-years prior in time $t-1$ (${}_5P_{10-5,t-1}$) to the population aged 0-4 in time t (${}_5P_{10-5,t}$). ie, if we have 100 5-9 year olds five years ago and we now have 125 10-14 year olds and 90 5-9 year olds, we can expect the number of 10-14 year olds in 5 years to be ($125/100 \cdot 90 = 112.5$).

Two age groups must have special consideration: the population aged 0-4 (${}_5P_0$) and the population comprising the open-ended interval (${}_\infty P_{85}$). The population aged 0-4 (${}_5P_0$) must have special consideration since the preceding/proceeding age groups do not exist for these age groups. To calculate the CCR for the open-ended age group,

$$\begin{aligned} {}_\infty CCR_{85,t} &= \frac{{}_\infty P_{85,t}}{{}_\infty P_{85-y,t-1}} \\ {}_\infty P_{85+t} &= {}_\infty CCR_{85,t} \cdot {}_\infty P_{85-y,t} \end{aligned} \tag{3}$$

Where y is the time difference between time periods.

For the population aged 0-4, we use the ratio of the population aged 0-4 to the number of women of reproductive age. Here we define women of reproductive age as the ages $[15, 50)$.

CCRs offer several advantages and disadvantages over the use of a cohort-component model. CCRs are considerably more parsimonious than cohort-component. Calculation of CCRs for use in population projections requires data as minimal as an age-sex distributions at two time periods – data ubiquitous across multiple scales, countries, and time periods. However, this parsimony comes at a relatively steep price: CCRs can lead to impossibly explosive growth in long-range projections due to the natural compounding of the ratios. Consider the growth currently occurring in McKenzie County, North Dakota (FIPS=38053) driving by the Shale oil boom. In 2010 McKenzie had a population of 6,360 that had ballooned to 12,792 by 2015, according to the Vintage 2016 population estimates from the US Census Bureau with a CCR for the 20-24 year old population of 2.46 (416 to 1,027). Implementing a 50-year population projection using that CCR would create a projected population that is approximately 8,000 times larger (2.46^{10}) – clearly an improbable population given the small, rural nature of its population.

Cohort Change Differences

The implementation of CCRs naturally implies a multiplicative model, typically utilizing leslie matrices. It is possible, however, to implement an **additive** model by using the *difference* in population rather than the *ratio* of population.

$$
\begin{aligned}
CCD_t &= {}_nP_{x,t} - {}_nP_{x-y,t-1} \\
{}_nP_{x+t} &= CCD_t + {}_nP_{x-y,t}
\end{aligned}
\tag{4}
$$

Where ${}_nP_{x,t}$ is the population aged x to $x+n$ in time t and ${}_nP_{x-y,t}$ is the population aged x to $x+n-y$ in time t where y refers to the time difference between time periods. These CCDs are calculated for each age group a, for each sex group s, for each race group r, in each time period t, in county c. Thus to find the population of ten to fourteen year olds (${}_5P_{10}$) in five years ($t+1$), we add the differene of the population aged 10-14 in time t (${}_5P_{10,t}$) to the population aged 5-9 five-years prior in time $t-1$ (${}_5P_{10-5,t-1}$) to the population aged 0-4 in time t (${}_5P_{10-5,t}$). ie, if we have 100 5-9 year olds five years ago and we now have 125 10-14 year olds and 90 5-9 year olds, we can expect the number of 10-14 year olds in 5 years to be (125-100 + 90 = 115).

Projecting CCRs and CCDs

It is unlikely that CCRs will remain unchanged over the projection horizon. To account for possible changes in CCRs, I employed the use of an unobserved components model (UCM) for forecasting equally spaced univariate time series data (Harvey 1990). UCMs decompose a time series into components such as trends, seasons, cycles, and regression effects and are designed to capture the features of the series that explain and predict its behavior. UCMs are similar to dynamic models in Bayesian time series forecasting (Harrison and West 1999). All projections were undertaken in R using the RUCM package.

The basic structural model (BSM) is the sum of its stochastic components. Here I use a trend component μ_t and a random error component ε_t and it can be described as:

$$y_t = \mu_t + \varepsilon_t \tag{5}$$

Each of the model components are modeled separately with the random error ε_t modeled as a sequence of independent, identically distributed zero-mean Gaussian random variables. The trend component is modeled using the following equations:

$$
\begin{aligned}
\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\
\beta_t &= \beta_{t-1} + \xi_t \\
\eta_t &\sim N(0, \sigma_\eta^2) \\
\xi_t &\sim N(0, \sigma_\xi^2)
\end{aligned}
$$

These equations specify a trend where the level μ_t and the slope β_t vary over time, governed by the variance of the disturbance terms η_t and ξ_t in their equations. Here all individual CCRs/CCDs (CCR_{iasr}) over the series were modelled (n=339,444) in individual UCM models.

Rather than use the prediction intervals output from the UCMs, I set the upper and lower bounds as the projected UCM plus or minus the 80th percentile based on the standard deviation of the original time series.

We forecast these UCMs for each CWR within a constrained forecast interval. CWRs are constrained to lie between (a, b). We limited CWRs such that each age/race/county combination would be constrained within the maximum/minimum of the time series such that $a = 0$ for all projections. and $b = max(CWR_{arc})$. We then transform the data using a scaled logit transformation to map (a, b) to the whole real line

$$
y = log(\frac{x - a}{b - x})
$$

Where x is the original data and y is the transformed data. The prediction intervals from these transformations have the same coverage probability as on the transformed scale, because quantiles are preserved under monotonically increasing transformations.

The projected CCRs and CCDs are then input into Leslie matrices to create projected populations:

$$
\begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{18} \end{bmatrix}_{t+1} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ CCR_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & CCR_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & CCR_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & CCR_{16} & CCR_{17} \end{bmatrix} \cdot \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{17} \end{bmatrix}_t
$$

$$
\mathbf{T} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ CCD_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & CCD_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & CCD_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & CCD_{16} & CCD_{17} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ n_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & n_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & n_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & n_{16} & n_{17} \end{bmatrix}
$$

$$
P_{t+1} \equiv \begin{bmatrix} \sum \mathbf{T_{1i}} \\ \sum \mathbf{T_{2i}} \\ \vdots \\ \sum \mathbf{T_{17i}} \end{bmatrix}
$$

The population aged 0-4 in time $t + 1$ are projected by applying a 1.05 sex ratio at birth (SRB) to the women of childbearing age $[15, 50)$ in time $t + 1$.

Extra considerations

These projections were carried out with 18 age groups (0,85,5), 2 sex groups, and 3 race groups (White, Black, Other).

All *resident* populations are projected in this modelling scheme such that the populations at launch year are equal to the total population minus the group quarters population. Group quarters populations at time t are then added back into the resident population at time $t + 1$.

Several county boundaries have also shifted since 1980:

- FIPS 12025 was changed to 12086.
- FIPS 15005 was absorbed by FIPS 15009.
- FIPS 51780 was merged into 51083.
- FIPS 51560 was merged into 51005.
- FIPS 30113 was split into 30031 and 30067. All three have been merged into 30031 – the larger county.
- FIPS 08014 was created out of parts of 08013, 08123, 08001, and 08059. Over 90% of the created population came out of 08013 so it is remerged.
- FIPS 02105 was created from 02105, 02230, and 02232 were all created out of the same 02230. 02230 was changed in 1992 from 02231.
- FIPS 02130, 02195, 02198, 02201, 02275, and 02280 were carved out of 02130.
- FIPS 02270 was recoded to 02158.
- FIPS 46113 was recoded to 46102.

In the event a UCM contained NA or infinite values or produced covariance matrices with values larger than 10,000,000, the projections were set to 0. Upper and Lower bounds of failed UCMs were set to 0. Additionally, any infinite, NA, or NAN CCR, CCD, or CWR was set to 0.

States included in this analysis: AL, AK, AZ, AR, CA, CO, CT, DE, DC, FL, GA, HI, ID, IL, IN, IA, KS, KY, LA, ME, MD, MA, MI, MN, MS, MO, MT, NE, NV, NH, NJ, NM, NY, NC, ND, OH, OK, OR, PA, RI, SC, SD, TN, TX, UT, VT, VA, WA, WV, WI, WY

Total number of counties: 3136

Overall Errors

Table 1 reports the overall errors for the sum of the population in each of the subsequent states and counties. Overall the purely ADDITIVE model outperformed the purely MULTIPLICATIVE model, suggesting CCDs could produce more accurate results compared to CCRs.

Table 1: Evaluation of TOTAL Errors. MAPE refers to MEDIAN Absolute Percent Error

TYPE	YEAR	POPULATION	PRED	LOW	HIGH	MAPE
ADD	2005	298,379,612	303,663,147	285,340,185	322,272,611	1.77%
ADD	2010	309,347,527	322,957,965	285,381,737	361,545,453	4.40%
ADD	2015	320,894,895	342,957,970	285,495,854	402,884,822	6.88%
ADDMULT	2005	298,373,465	304,674,178	286,505,559	323,211,601	2.11%
ADDMULT	2010	309,344,478	326,521,302	289,545,501	367,786,796	5.55%
ADDMULT	2015	320,890,305	353,678,081	292,224,571	444,354,649	10.22%
Mult	2005	298,379,612	310,264,056	288,517,257	332,581,349	4.0%
Mult	2010	309,341,025	348,226,543	298,099,857	411,654,046	12.6%
Mult	2015	320,885,858	421,672,155	313,302,023	630,712,200	31.4%

```
## Joining, by = "STATEID"
```

Table 2: Evaluation of STATE Errors. MAPE refers to MEDIAN Absolute Percent Error

TYPE	COUNTYnum	state	2005	2010	2015
ADD	25	AK	5.00%	9.96%	11.03%
ADDMULT	25	AK	3.76%	7.23%	10.65%
Mult	25	AK	3.75%	8.47%	12.45%
ADD	67	AL	3.261%	5.79%	8.69%
ADDMULT	67	AL	3.412%	5.79%	10.09%
Mult	67	AL	4.176%	7.51%	18.26%
ADD	75	AR	2.37%	5.69%	10.5%
ADDMULT	75	AR	2.92%	7.03%	14.1%
Mult	75	AR	4.94%	13.34%	36.1%
ADD	15	AZ	2.520%	11.624%	19.902%
ADDMULT	15	AZ	2.520%	11.624%	19.902%
Mult	15	AZ	2.806%	12.568%	20.476%
ADD	58	CA	3.975%	7.52%	12.00%
ADDMULT	58	CA	3.891%	7.12%	12.00%
Mult	58	CA	4.690%	10.61%	18.07%
ADD	63	CO	7.515%	14.50%	20.6%
ADDMULT	63	CO	7.515%	14.50%	21.4%
Mult	63	CO	8.129%	19.11%	40.6%
ADD	8	CT	1.353%	2.46%	4.82%
ADDMULT	8	CT	1.353%	2.46%	4.82%
Mult	8	CT	1.141%	4.99%	13.67%
ADD	1	DC	1.26%	9.01%	20.0%
ADDMULT	1	DC	0.20%	2.68%	5.9%
Mult	1	DC	0.20%	2.68%	5.9%
ADD	3	DE	2.54%	8.33%	11.2%
ADDMULT	3	DE	2.54%	8.33%	11.2%
Mult	3	DE	6.04%	11.23%	22.0%
ADD	67	FL	3.174%	5.64%	10.1%
ADDMULT	67	FL	3.174%	5.64%	9.4%
Mult	67	FL	3.740%	14.35%	33.5%
ADD	159	GA	4.10%	8.46%	16.1%
ADDMULT	159	GA	4.06%	8.46%	16.3%
Mult	159	GA	5.78%	15.65%	34.9%
ADD	4	HI	4.99%	7.11%	8.51%
ADDMULT	4	HI	4.99%	7.11%	8.51%
Mult	4	HI	3.31%	3.09%	3.63%
ADD	99	IA	1.480%	3.18%	5.60%
ADDMULT	99	IA	1.406%	3.52%	6.54%
Mult	99	IA	1.584%	4.25%	7.49%
ADD	44	ID	4.542%	9.09%	15.23%
ADDMULT	44	ID	4.431%	9.36%	15.23%
Mult	44	ID	4.945%	13.28%	20.95%
ADD	102	IL	1.631%	3.45%	4.8%
ADDMULT	102	IL	2.220%	5.02%	12.7%
Mult	102	IL	2.243%	6.46%	17.0%
ADD	92	IN	2.04%	4.59%	7.0%
ADDMULT	92	IN	2.33%	5.43%	9.5%
Mult	92	IN	3.42%	9.53%	27.0%
ADD	105	KS	2.575%	6.832%	10.15%
ADDMULT	105	KS	3.030%	6.976%	11.94%

TYPE	COUNTYnum	state	2005	2010	2015
Mult	105	KS	3.504%	7.758%	16.78%
ADD	120	KY	2.708%	6.04%	10.2%
ADDMULT	120	KY	2.483%	6.13%	10.2%
Mult	120	KY	3.348%	9.94%	20.9%
ADD	64	LA	2.554%	4.95%	6.76%
ADDMULT	64	LA	2.299%	3.67%	5.78%
Mult	64	LA	2.431%	4.96%	13.59%
ADD	14	MA	3.12%	4.96%	6.2%
ADDMULT	14	MA	3.12%	5.45%	9.5%
Mult	14	MA	4.73%	9.33%	16.9%
ADD	24	MD	2.569%	5.72%	6.8%
ADDMULT	24	MD	3.001%	5.89%	9.4%
Mult	24	MD	3.062%	11.76%	27.9%
ADD	16	ME	3.17%	6.36%	9.5%
ADDMULT	16	ME	2.52%	5.95%	8.7%
Mult	16	ME	4.27%	7.45%	20.1%
ADD	83	MI	3.65%	9.10%	14.0%
ADDMULT	83	MI	4.09%	9.64%	17.6%
Mult	83	MI	5.10%	16.36%	33.3%
ADD	87	MN	2.08%	5.89%	10.45%
ADDMULT	87	MN	2.53%	5.89%	11.53%
Mult	87	MN	3.11%	8.77%	17.81%
ADD	115	MO	2.36%	6.01%	11.0%
ADDMULT	115	MO	2.45%	5.69%	11.9%
Mult	115	MO	4.12%	12.48%	32.3%
ADD	82	MS	2.82%	5.30%	9.3%
ADDMULT	82	MS	3.56%	6.76%	14.0%
Mult	82	MS	4.71%	9.19%	19.6%
ADD	55	MT	3.983%	8.97%	15.13%
ADDMULT	55	MT	3.249%	7.21%	12.55%
Mult	55	MT	3.239%	7.87%	13.05%
ADD	100	NC	2.96%	6.59%	12.5%
ADDMULT	100	NC	2.88%	6.07%	12.8%
Mult	100	NC	4.66%	14.45%	33.2%
ADD	53	ND	3.383%	11.47%	24.89%
ADDMULT	53	ND	2.464%	8.31%	18.67%
Mult	53	ND	2.896%	8.31%	18.67%
ADD	93	NE	2.337%	6.318%	9.18%
ADDMULT	93	NE	2.052%	5.344%	8.15%
Mult	93	NE	2.537%	6.151%	10.88%
ADD	10	NH	3.47%	8.36%	13.8%
ADDMULT	10	NH	2.58%	8.36%	19.1%
Mult	10	NH	4.27%	15.02%	33.1%
ADD	21	NJ	3.538%	7.21%	9.91%
ADDMULT	21	NJ	2.984%	5.66%	9.91%
Mult	21	NJ	3.823%	8.81%	18.12%
ADD	33	NM	6.391%	14.57%	23.85%
ADDMULT	33	NM	6.391%	13.74%	19.21%
Mult	33	NM	6.511%	14.89%	19.84%
ADD	17	NV	7.19%	18.98%	32.07%
ADDMULT	17	NV	7.19%	16.47%	29.17%
Mult	17	NV	8.53%	22.87%	34.20%

TYPE	COUNTYnum	state	2005	2010	2015
ADD	62	NY	2.301%	4.66%	6.26%
ADDMULT	62	NY	1.481%	2.61%	5.37%
Mult	62	NY	1.652%	4.35%	7.05%
ADD	88	OH	1.34%	3.35%	5.6%
ADDMULT	88	OH	1.81%	5.74%	13.6%
Mult	88	OH	3.18%	12.23%	39.8%
ADD	77	OK	2.966%	5.92%	8.76%
ADDMULT	77	OK	2.725%	3.90%	7.09%
Mult	77	OK	3.344%	7.53%	15.97%
ADD	36	OR	3.170%	6.50%	10.05%
ADDMULT	36	OR	2.863%	5.02%	9.56%
Mult	36	OR	3.182%	7.85%	16.74%
ADD	67	PA	1.976%	4.351%	6.52%
ADDMULT	67	PA	1.793%	3.695%	8.18%
Mult	67	PA	1.815%	4.586%	14.78%
ADD	5	RI	2.74%	7.10%	9.4%
ADDMULT	5	RI	2.74%	7.10%	9.4%
Mult	5	RI	4.87%	13.67%	26.4%
ADD	46	SC	2.65%	6.00%	10.9%
ADDMULT	46	SC	2.65%	6.13%	11.8%
Mult	46	SC	4.36%	11.40%	26.9%
ADD	66	SD	3.479%	5.54%	11.07%
ADDMULT	66	SD	2.973%	5.82%	9.95%
Mult	66	SD	3.273%	7.46%	12.06%
ADD	95	TN	3.12%	7.66%	13.1%
ADDMULT	95	TN	3.20%	7.69%	13.4%
Mult	95	TN	5.21%	15.56%	32.8%
ADD	254	TX	3.439%	7.55%	13.08%
ADDMULT	254	TX	3.287%	7.23%	13.10%
Mult	254	TX	4.020%	10.78%	21.38%
ADD	29	UT	3.59%	6.94%	11.25%
ADDMULT	29	UT	3.58%	6.75%	10.61%
Mult	29	UT	4.69%	7.91%	16.06%
ADD	133	VA	NA	6.00%	10.2%
ADD	134	VA	3.086%	NA	NA
ADDMULT	133	VA	3.026%	5.95%	9.9%
Mult	133	VA	NA	11.53%	27.2%
Mult	134	VA	3.874%	NA	NA
ADD	14	VT	2.07%	3.83%	5.8%
ADDMULT	14	VT	2.33%	5.26%	9.1%
Mult	14	VT	3.40%	11.03%	30.7%
ADD	39	WA	2.224%	5.16%	6.87%
ADDMULT	39	WA	2.224%	5.16%	6.87%
Mult	39	WA	2.953%	6.31%	13.80%
ADD	72	WI	1.43%	4.22%	7.92%
ADDMULT	72	WI	1.52%	4.37%	8.20%
Mult	72	WI	2.43%	7.23%	16.44%
ADD	55	WV	3.28%	7.34%	11.96%
ADDMULT	55	WV	1.29%	4.01%	12.72%
Mult	55	WV	1.55%	5.77%	14.53%
ADD	23	WY	4.67%	11.11%	14.37%
ADDMULT	23	WY	4.10%	10.66%	10.70%

TYPE	COUNTYnum	state	2005	2010	2015
Mult	23	WY	3.40%	8.92%	8.74%

The total error for any given county is also small and only marginally larger than the nationwide total.

Table 3: Evaluation of TOTAL Errors for counties. MAPE refers to MEDIAN Absolute Percent Error

COUNTYnum	TYPE	VAR	2005	2010	2015
3135	ADD	MAPE	NA	6.17%	10.4%
3136	ADD	MAPE	2.780%	NA	NA
3135	Mult	MAPE	NA	9.55%	20.6%
3136	Mult	MAPE	3.553%	NA	NA
3135	ADDMULT	MAPE	2.767%	6.18%	11.4%
3135	ADD	in 80th percentile	NA	89.92%	87.4%
3136	ADD	in 80th percentile	91.87%	NA	NA
3135	Mult	in 80th percentile	NA	83.51%	77.2%
3136	Mult	in 80th percentile	89.99%	NA	NA
3135	ADDMULT	in 80th percentile	92.03%	89.76%	86.5%

Errors by Age

The errors for age groups are also relatively low with the average age group having an overall error of 13%.

Table 4: Evaluation of Age Group Errors. MAPE refers to MEDIAN Absolute Percent Error

num	TYPE	VAR	2005	2010	2015
56430	ADD	MAPE	NA	0.0946	0.1399
56448	ADD	MAPE	0.0546	NA	NA
56430	Mult	MAPE	NA	0.1120	0.1809
56448	Mult	MAPE	0.0613	NA	NA
56430	ADD	in 80th percentile	NA	0.7181	0.7393
56448	ADD	in 80th percentile	0.6353	NA	NA
56430	Mult	in 80th percentile	NA	0.6985	0.7197
56448	Mult	in 80th percentile	0.6164	NA	NA

Errors by Sex

Table 5: Evaluation of Sex Errors. MAPE refers to MEDIAN Absolute Percent Error

num	SEX	TYPE	YEAR	MAPE	in80percentile
3136	FEMALE	ADD	2005	2.642%	92.41%
3135	FEMALE	ADDMULT	2005	2.548%	92.41%
3136	FEMALE	Mult	2005	3.281%	90.69%
3136	MALE	ADD	2005	3.043%	89.92%
3135	MALE	ADDMULT	2005	3.035%	90.21%

num	SEX	TYPE	YEAR	MAPE	in80percentile
3136	MALE	Mult	2005	3.902%	88.30%
3135	FEMALE	ADD	2010	5.84%	90.81%
3135	FEMALE	ADDMULT	2010	5.65%	90.94%
3135	FEMALE	Mult	2010	8.79%	84.75%
3135	MALE	ADD	2010	6.68%	87.88%
3135	MALE	ADDMULT	2010	6.78%	87.59%
3135	MALE	Mult	2010	10.25%	81.88%
3135	FEMALE	ADD	2015	9.76%	88.5%
3135	FEMALE	ADDMULT	2015	10.48%	87.8%
3135	FEMALE	Mult	2015	18.63%	78.0%
3135	MALE	ADD	2015	11.0%	85.68%
3135	MALE	ADDMULT	2015	12.2%	84.37%
3135	MALE	Mult	2015	21.1%	76.20%

Errors by Race

Table 6: Evaluation of Race Errors. MAPE refers to MEDIAN Absolute Percent Error

num	RACE	TYPE	YEAR	MAPE	in80percentile
3136	BLACK	ADD	2005	10.48%	76.08%
3135	BLACK	ADD	2010	17.31%	77.89%
3135	BLACK	ADD	2015	22.64%	81.75%
3135	BLACK	ADDMULT	2005	10.93%	70.97%
2931	BLACK	ADDMULT	2010	16.73%	76.36%
2931	BLACK	ADDMULT	2015	22.47%	79.05%
3136	BLACK	Mult	2005	12.78%	67.98%
2847	BLACK	Mult	2010	19.34%	75.13%
2847	BLACK	Mult	2015	27.90%	74.68%
3136	OTHER	ADD	2005	25.7%	38.6%
3135	OTHER	ADD	2010	53.5%	29.98%
3135	OTHER	ADD	2015	70%	31.90%
3135	OTHER	ADDMULT	2005	29.9%	36.7%
3114	OTHER	ADDMULT	2010	62.5%	30.03%
3114	OTHER	ADDMULT	2015	83%	30.22%
3136	OTHER	Mult	2005	64.6%	26.9%
3109	OTHER	Mult	2010	145.8%	24.96%
3109	OTHER	Mult	2015	279%	25.41%
3136	WHITE	ADD	2005	2.865%	89.349%
3135	WHITE	ADD	2010	6.613%	86.252%
3135	WHITE	ADD	2015	10.86%	83.764%
3135	WHITE	ADDMULT	2005	2.639%	89.442%
3135	WHITE	ADDMULT	2010	5.889%	86.507%
3135	WHITE	ADDMULT	2015	9.73%	83.477%
3136	WHITE	Mult	2005	2.850%	88.680%
3135	WHITE	Mult	2010	6.282%	85.933%
3135	WHITE	Mult	2015	10.20%	83.700%

Errors for all joint combinations

Table 7: Evaluation of Age/Sex/Race Errors. MAPE refers to MEDIAN Absolute Percent Error

num	TYPE	YEAR	MAPE	in80percentile
338060	ADD	2005	13.69%	44.242%
337952	ADDMULT	2005	13.44%	44.056%
338060	Mult	2005	15.26%	44.825%
337324	ADD	2010	22.20%	52.89%
330124	ADDMULT	2010	21.71%	51.80%
327276	Mult	2010	25.54%	52.38%
336696	ADD	2015	31.23%	55.92%
329946	ADDMULT	2015	30.52%	54.20%
327276	Mult	2015	38.08%	54.47%