

# Suppressed Data and Imputation Appendix

to

Rust and Renewal: A Cincinnati Retrospective

Rust and Renewal: A Cleveland Retrospective

## Industrial Heartland Series

February 2018

FEDERAL RESERVE BANK *of* CLEVELAND

## Suppressed Data and Imputation Appendix

This appendix accompanies the reports “Rust and Recovery: A Cincinnati Retrospective” and “Rust and Recovery: A Cleveland Retrospective.” Suppressed data were imputed for some of the charts in the reports, and this appendix documents the suppressed observations in each MSA and outlines the imputation method. It also provides an assessment of the imputations. The Bureau of Economic Analysis will suppress data for metropolitan areas when doing so is necessary to avoid identification of individual establishments.

Suppressed data were imputed for three figures in the Cincinnati MSA report and one figure in the Cleveland MSA report.

- Cincinnati MSA figures with imputed data:
  - Figure 2: Cincinnati MSA Sector Shares of Employment (26 imputed observations)
  - Figure 4: Manufacturing Share of Total Employment (1 imputed observation, manufacturing employment in 2015)
  - Figure 5: Manufacturing Share of Earnings (1 imputed observation, manufacturing earnings in 2015)
- Cleveland MSA figure with imputed data:
  - Figure 3: Cleveland MSA Sector Shares of Employment (9 imputed observations)

## Employment Composition Charts

Sector-level data are classified according to the Standard Industrial Codes (SIC) for 1969–2000 and by the North American Industry Classification System (NAICS) for 2001–2016. We used the following SIC–NAICS crosswalk to generate the sectors identified in the following table:

Sector	SIC line code from data file CA25	NAICS line code from data file CA25N
Construction	300	400
Manufacturing	400	500
Transportation and public utilities	500	300, 800
Wholesale trade	610	600
Retail trade	620	700
Finance, insurance, and real estate	700	1000, 1100
Services	800	900, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900
Total employment	10	10
<b>Total number of observations</b>	<b>256 (= 32 years x 8 line codes)</b>	<b>288 (= 16 years x 18 line codes)</b>

- The employment composition charts in the Cincinnati and Cleveland reports, figures 2 and 3, respectively, have suppressed data.
- Suppressed data were imputed at the individual line-code level for both SIC and NAICS before being aggregated to create the sector values.

## Imputation Methodology and Assessment

If three or more observations are present for a given SIC or NAICS line code, any suppressed values are imputed using the Kalman smoother available from the “imputeTS” R package.

If fewer than three observations are present for a given SIC or NAICS line code, any suppressed values are imputed using the mean of all observed data.

The imputation methodology is assessed as follows:

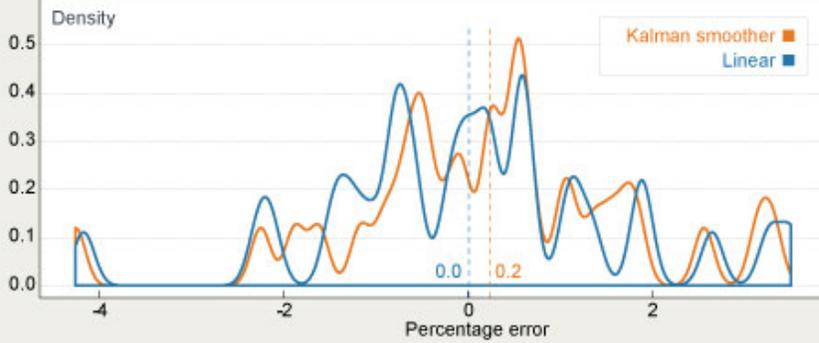
- Full sample size is 544 observations (256 SIC observations and 288 NAICS observations)
- Randomly insert, with replacement, 5 percent, 10 percent, and 15 percent missing values into the 544 observations
- Impute the missing values using (a) the approach described above and (b) simple linear interpolation
- Compute the percentage error between the actual value and the imputed value
- Repeat the process 5,000 times and plot the density function of errors for a given percentage of missing observations
- The median imputation error is always less than or equal to 0.4 percent

## Suppressed Employment Data for the Cincinnati MSA

Sector	GeoFIPS	Description	Number suppressed	Observations per line code	Share suppressed (%)
NAICS 10	17140	Total employment (number of jobs)	0	16	0
NAICS 1000	17140	Finance and insurance	0	16	0
NAICS 1100	17140	Real estate and rental and leasing	1	16	6.25
NAICS 1200	17140	Professional, scientific, and technical services	3	16	18.75
NAICS 1300	17140	Management of companies and enterprises	6	16	37.50
NAICS 1400	17140	Administrative and support and waste management and remediation services	0	16	0
NAICS 1500	17140	Educational services	0	16	0
NAICS 1600	17140	Health care and social assistance	0	16	0
NAICS 1700	17140	Arts, entertainment, and recreation	4	16	25.00
NAICS 1800	17140	Accommodation and food services	0	16	0
NAICS 1900	17140	Other services (except public administration)	2	16	12.50
NAICS 300	17140	Utilities	7	16	43.75
NAICS 400	17140	Construction	0	16	0
NAICS 500	17140	Manufacturing	1	16	6.25
NAICS 600	17140	Wholesale trade	1	16	6.25
NAICS 700	17140	Retail trade	0	16	0
NAICS 800	17140	Transportation and warehousing	0	16	0
NAICS 900	17140	Information	0	16	0
SIC 10	17140	Total employment (number of jobs)	0	32	0
SIC 300	17140	Construction	1	32	3.13
SIC 400	17140	Manufacturing	0	32	0
SIC 500	17140	Transportation and public utilities	0	32	0
SIC 610	17140	Wholesale trade	0	32	0
SIC 620	17140	Retail trade	0	32	0
SIC 700	17140	Finance, insurance, and real estate	0	32	0
SIC 800	17140	Services	0	32	0
<b>TOTAL</b>			<b>26</b>	<b>544</b>	<b>4.78</b>

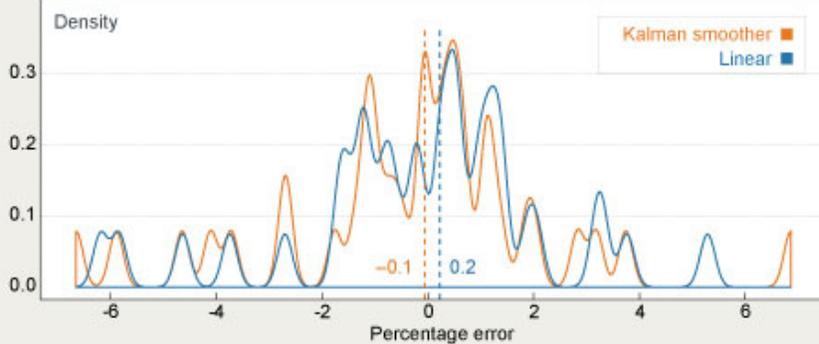
Calculations from the Bureau of Economic Analysis data files CA25 and CA25N.

### Density Plots of Employment Imputation Errors: Cincinnati MSA



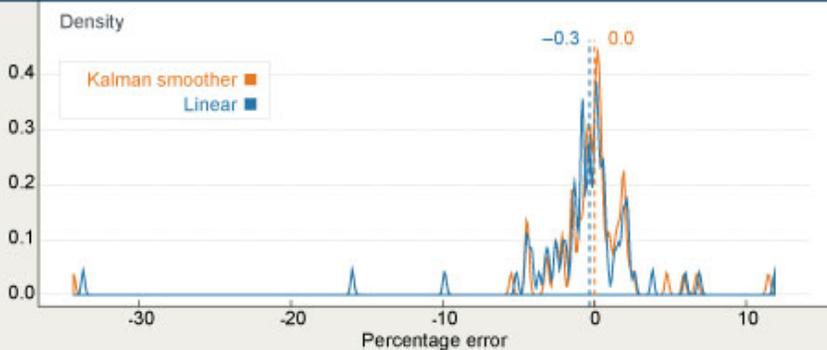
Density plots are based on 5,000 replications with 5% missing values randomly inserted. Median values of each density plot are illustrated as dashed vertical lines.

### Density Plots of Employment Imputation Errors: Cincinnati MSA



Density plots are based on 5,000 replications with 10% missing values randomly inserted. Median values of each density plot are illustrated as dashed vertical lines.

### Density Plots of Employment Imputation Errors: Cincinnati MSA



Density plots are based on 5,000 replications with 15% missing values randomly inserted. Median values of each density plot are illustrated as dashed vertical lines.

## Suppressed Employment Data for the Cleveland MSA

Sector	GeoFIPS	Description	Number suppressed	Observations per line code	Share suppressed (%)
NAICS 10	17460	Total employment (number of jobs)	0	16	0
NAICS 1000	17460	Finance and insurance	0	16	0
NAICS 1100	17460	Real estate and rental and leasing	0	16	0
NAICS 1200	17460	Professional, scientific, and technical services	0	16	0
NAICS 1300	17460	Management of companies and enterprises	0	16	0
NAICS 1400	17460	Administrative and support and waste management and remediation services	0	16	0
NAICS 1500	17460	Educational services	0	16	0
NAICS 1600	17460	Health care and social assistance	0	16	0
NAICS 1700	17460	Arts, entertainment, and recreation	0	16	0
NAICS 1800	17460	Accommodation and food services	0	16	0
NAICS 1900	17460	Other services (except public administration)	0	16	0
NAICS 300	17460	Utilities	1	16	6.25
NAICS 400	17460	Construction	0	16	0
NAICS 500	17460	Manufacturing	0	16	0
NAICS 600	17460	Wholesale trade	0	16	0
NAICS 700	17460	Retail trade	0	16	0
NAICS 800	17460	Transportation and warehousing	8	16	50
NAICS 900	17460	Information	0	16	0
SIC 10	17460	Total employment (number of jobs)	0	32	0
SIC 300	17460	Construction	0	32	0
SIC 400	17460	Manufacturing	0	32	0
SIC 500	17460	Transportation and public utilities	0	32	0
SIC 610	17460	Wholesale trade	0	32	0
SIC 620	17460	Retail trade	0	32	0
SIC 700	17460	Finance, insurance, and real estate	0	32	0
SIC 800	17460	Services	0	32	0
<b>TOTAL</b>			<b>9</b>	<b>544</b>	<b>1.65</b>

Calculations from the Bureau of Economic Analysis data files CA25 and CA25N.

