

```

1 library(readr)
2 library(dplyr)
3 library(reshape2)
4 library(survey)
5
6 # Export 30 from IPUMS
7 # "Method of travel to work, wages, occupation, industry, class of worker, and usual
  hours worked (1980, 1990, 2000-2015) w/ CSV"
8
9 # Read IPUMS export
10 ipums.orig <- read_csv("usa_00030.csv", col_types="ciiiiiidi")
11
12 # Filter to only full-time wageworkers (>= 35 hours per week, with wages, and not self-
  employed)
13 ipums <- ipums.orig %>%
14   filter(INCWAGE > 0 & INCWAGE < 999998 & CLASSWKR == 2 & UHRSWORK >= 35)
15
16 # Convert years to a factor
17 ipums$year <- as.factor(ipums$YEAR)
18
19 # Convert TRANWORK
20 ipums$commute <- NA
21
22 ipums$commute[ipums$TRANWORK >= 10 & ipums$TRANWORK <= 20] <- "Private vehicle"
23 ipums$commute[ipums$TRANWORK >= 30 & ipums$TRANWORK <= 36] <- "Public transit"
24 ipums$commute[ipums$TRANWORK == 40] <- "Bicycle"
25 ipums$commute[ipums$TRANWORK == 50] <- "Walked"
26 ipums$commute[ipums$TRANWORK == 60] <- "Other"
27 ipums$commute[ipums$TRANWORK == 70] <- "Worked at home"
28
29 ipums$commute <- factor(ipums$commute, level=c("Private vehicle", "Public transit",
  "Bicycle", "Walked", "Other", "Worked at home"))
30
31 # Recode OCC2010 to match ATUS OCC2 coding
32 ipums$jobs <- NA
33
34 ipums$jobs[ipums$OCC2010 >= 10 & ipums$OCC2010 <= 430] <- "Management occupations"
35 ipums$jobs[ipums$OCC2010 >= 500 & ipums$OCC2010 <= 950] <- "Business and financial
  operations occupations"
36 ipums$jobs[ipums$OCC2010 >= 1000 & ipums$OCC2010 <= 1240] <- "Computer and mathematical
  science occupations"
37 ipums$jobs[ipums$OCC2010 >= 1300 & ipums$OCC2010 <= 1560] <- "Architecture and
  engineering occupations"
38 ipums$jobs[ipums$OCC2010 >= 1600 & ipums$OCC2010 <= 1980] <- "Life, physical, and social
  science occupations"
39 ipums$jobs[ipums$OCC2010 >= 2000 & ipums$OCC2010 <= 2060] <- "Community and social
  service occupations"
40 ipums$jobs[ipums$OCC2010 >= 2100 & ipums$OCC2010 <= 2150] <- "Legal occupations"
41 ipums$jobs[ipums$OCC2010 >= 2200 & ipums$OCC2010 <= 2550] <- "Education, training, and
  library occupations"
42 ipums$jobs[ipums$OCC2010 >= 2600 & ipums$OCC2010 <= 2920] <- "Arts, design,
  entertainment, sports, and media occupations"
43 ipums$jobs[ipums$OCC2010 >= 3000 & ipums$OCC2010 <= 3540] <- "Healthcare practitioner and
  technical occupations"
44 ipums$jobs[ipums$OCC2010 >= 3600 & ipums$OCC2010 <= 3650] <- "Healthcare support
  occupations"

```

```

45 ipums$jobs[ipums$OCC2010 >= 3700 & ipums$OCC2010 <= 3950] <- "Protective service
    occupations"
46 ipums$jobs[ipums$OCC2010 >= 4000 & ipums$OCC2010 <= 4150] <- "Food preparation and
    serving related occupations"
47 ipums$jobs[ipums$OCC2010 >= 4200 & ipums$OCC2010 <= 4250] <- "Building and grounds
    cleaning and maintenance occupations"
48 ipums$jobs[ipums$OCC2010 >= 4300 & ipums$OCC2010 <= 4650] <- "Personal care and service
    occupations"
49 ipums$jobs[ipums$OCC2010 >= 4700 & ipums$OCC2010 <= 4965] <- "Sales and related
    occupations"
50 ipums$jobs[ipums$OCC2010 >= 5000 & ipums$OCC2010 <= 5940] <- "Office and administrative
    support occupations"
51 ipums$jobs[ipums$OCC2010 >= 6005 & ipums$OCC2010 <= 6130] <- "Farming, fishing, and
    forestry occupations"
52 ipums$jobs[ipums$OCC2010 >= 6200 & ipums$OCC2010 <= 6940] <- "Construction and extraction
    occupations"
53 ipums$jobs[ipums$OCC2010 >= 7000 & ipums$OCC2010 <= 7630] <- "Installation, maintenance,
    and repair occupations"
54 ipums$jobs[ipums$OCC2010 >= 7700 & ipums$OCC2010 <= 8965] <- "Production occupations"
55 ipums$jobs[ipums$OCC2010 >= 9000 & ipums$OCC2010 <= 9750] <- "Transportation and material
    moving occupations"
56
57 ipums$jobs <- as.factor(ipums$jobs)
58
59 # Read occupation codes
60 occ <-
    read_csv("https://raw.githubusercontent.com/wireservice/lookup/master/occ/description.201
    0.csv", col_types="ic")
61
62 occ$description <- factor(occ$description)
63
64 # Read industry codes
65 ind <- read_csv("ind1990.csv", col_types="ic")
66
67 ind$description <- factor(ind$description)
68
69 # Read CPI rates
70 cpi <-
    read_csv("https://raw.githubusercontent.com/wireservice/lookup/master/year/cpi.csv",
    col_types="cd")
71
72 # Function to perform inflation adjustment on a list of columns
73 cpi.adjust <- function(input, year.column, columns) {
74   temp <- left_join(input, cpi, by = year.column)
75
76   for (column in columns) {
77     temp[[column]] <- temp[[column]] * 237.0 / temp[["cpi"]]
78   }
79
80   temp$cpi <- NULL
81
82   return(temp)
83 }
84
85 # Compute mean wages by year and commute
86 # NB: Results are a perfect match for IPUMS online analysis
87 ipums.commute.means <- ipums %>%

```

```

88   group_by(year, commute) %>%
89   summarise(
90     pop = sum(PERWT),
91     mean.wages = sum(INCWAGE * PERWT) / pop,
92     v = (sum(PERWT * (INCWAGE - mean.wages)^2)) / (pop - 1),
93     sd = sqrt(v),
94     se = sd / sqrt(pop)
95   )
96
97   ipums.commute.means.cpi <- cpi.adjust(ipums.means, "year", c("mean.wages", "v", "sd",
98     "se"))
99
100  write_csv(ipums.commute.means.cpi, "results/ipums.commute.means.cpi.csv")
101
102  # Compute mean wages for home workers by year and occupation
103  ipums.occ.means <- ipums %>%
104    filter(commute == "Worked at home") %>%
105    group_by(year, OCC2010) %>%
106    summarise(
107      pop = sum(PERWT),
108      mean.wages = sum(INCWAGE * PERWT) / pop,
109      v = (sum(PERWT * (INCWAGE - mean.wages)^2)) / (pop - 1),
110      sd = sqrt(v),
111      se = sd / sqrt(pop)
112    ) %>%
113    left_join(occ, by = c("OCC2010" = "occ"))
114
115  ipums.occ.means.cpi <- cpi.adjust(ipums.occ.means, "year", c("mean.wages", "v", "sd",
116    "se"))
117
118  write_csv(ipums.occ.means.cpi, "results/ipums.occ.means.cpi.csv")
119
120  # Compute mean wages for home workers by year and job category
121  ipums.jobs.means <- ipums %>%
122    filter(commute == "Worked at home") %>%
123    group_by(year, jobs) %>%
124    summarise(
125      pop = sum(PERWT),
126      mean.wages = sum(INCWAGE * PERWT) / pop,
127      v = (sum(PERWT * (INCWAGE - mean.wages)^2)) / (pop - 1),
128      sd = sqrt(v),
129      se = sd / sqrt(pop)
130    )
131
132  ipums.jobs.means.cpi <- cpi.adjust(ipums.jobs.means.cpi, "year", c("mean.wages", "v",
133    "sd", "se"))
134
135  write_csv(ipums.jobs.means.cpi, "results/ipums.jobs.means.cpi.csv")
136
137  # Compute mean wages by job category and commute type for 2015 only
138  ipums.jobs.2015.means <- ipums %>%
139    filter(year == 2015) %>%
140    group_by(jobs, commute) %>%
141    summarise(
142      pop = sum(PERWT),
143      mean.wages = sum(INCWAGE * PERWT) / pop,
144      v = (sum(PERWT * (INCWAGE - mean.wages)^2)) / (pop - 1),

```

```
142     sd = sqrt(v),
143     se = sd / sqrt(pop)
144   )
145
146 write_csv(ipums.jobs.2015.means, "results/ipums.jobs.2015.means.csv")
147
148 # Compute fraction of occupation that works from home by year
149 ipums.jobs.shares <- ipums %>%
150   group_by(year, jobs, commute) %>%
151   summarise(
152     pop = sum(PERWT)
153   ) %>%
154   mutate(
155     share = pop / sum(pop)
156   )
157
158 write_csv(ipums.jobs.shares, "results/ipums.jobs.shares.csv")
159
160 # Compute share of homeworkers by year and industry
161 ipums.ind.shares <- ipums %>%
162   group_by(year, IND1990, commute) %>%
163   summarise(
164     pop = sum(PERWT)
165   ) %>%
166   mutate(
167     share = pop / sum(pop)
168   ) %>%
169   left_join(ind, by = c("IND1990" = "ind"))
170
171 write_csv(ipums.ind.shares, "results/ipums.ind.shares.csv")
172
```