Convergence figure 5

Ari Anisfeld

March 6, 2019

```
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
library(yaml)
library(rprojroot)
library(broom)

make_path <- is_git_root$make_fix_file()
working_path <- make_path("journalist/eduardo_porter/")
CONFIG <- yaml.load_file(file.path(working_path, "/code/config.yml"))
source(str_c(CONFIG$lab_code, "prelim.R"))
source(file.path(working_path, "/code/binscatter.R"))
```

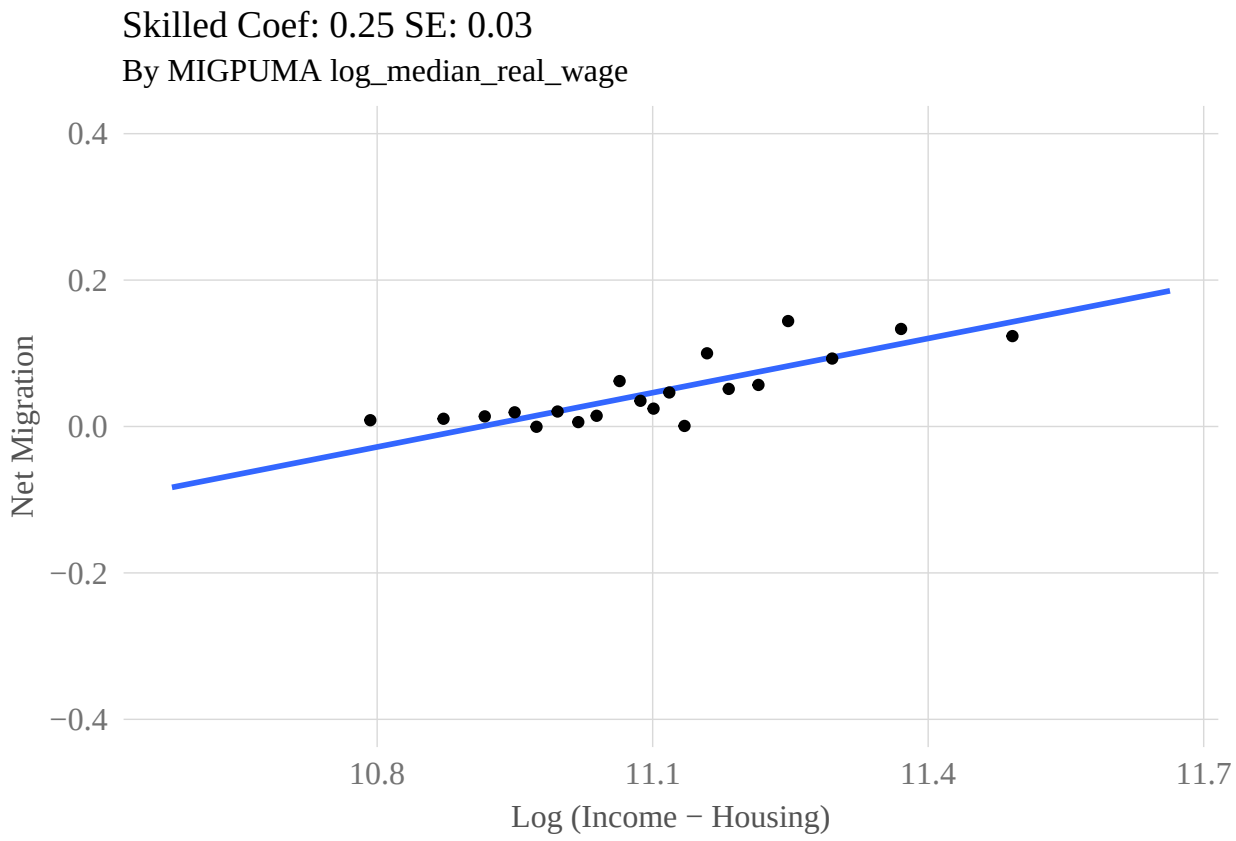
```
prepare_data <- function(data, name, place_type){
  data %>%
    # remove foreign places
    filter(!is.na(`net migration`)) %>%
    select(starts_with(place_type), starts_with("state"), `net migration`,
           starts_with("median"), starts_with("mean")) %>%
    mutate(skill = name,
           net_migration = `net migration`*100,
           log_median_real_wage = log(median_real_wage),
           log_median_nominal_wage = log(median_nominal_wage),
           log_mean_real_wage = log(mean_real_wage),
           log_mean_nominal_wage = log(mean_nominal_wage)
           )
}

get_model_as_title <- function(tidy_model, group="", round_to=2) {

    stopifnot( nrow(tidy_model) == 2)
    tidy_model <- tidy_model %>%
                    filter(term != "(Intercept)") %>%
                    transmute(Coef = estimate,
                              SE = std.error) %>%
                    round(round_to)


    title = group
    for(col in names(tidy_model)){
      title = glue::glue("{title} {col}: {tidy_model[, col]}")
    }
  title
}

make_plot <- function(data=migpuma_plot,
                      skill.="Skilled",
                      wage_type="log_median_real_wage",
```

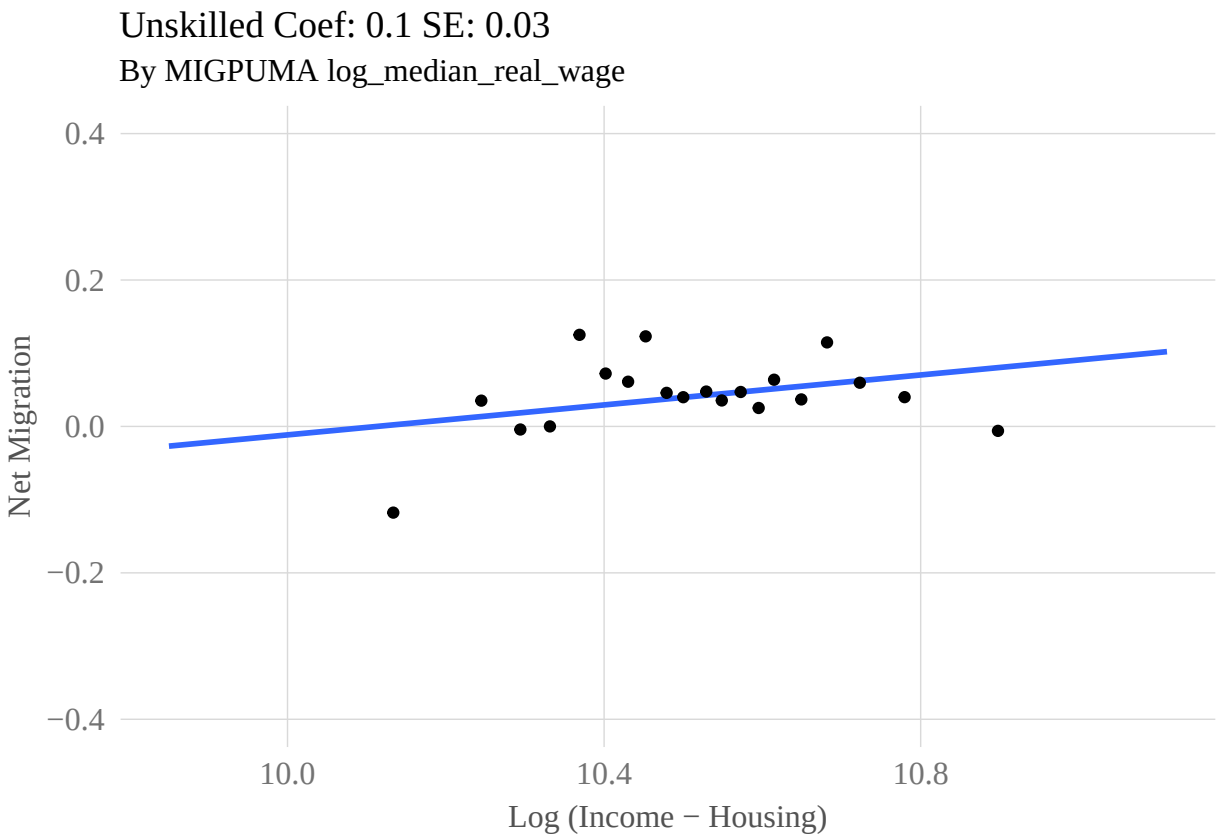
```
                      weights. = "migpuma_population",
                      ylims = .4){


  filtered_data <- data %>% filter(skill==skill.)


  binscatter_output <-
    filtered_data %>%
      binscatter(df = .,
                 x = wage_type,
                 y = "net_migration",
                 group = "skill",
                 weights = weights.
    )

  model <- lm(as.formula(str_c("net_migration~", wage_type)),
              data = filtered_data,
              weights = filtered_data %>% pull(!!sym(weights.)))
  tidy_model <- tidy(model)

  x_label <- ifelse(str_detect(wage_type, "real"), "Log (Income - Housing)", "Log (Income)")

  filtered_data %>%
    ggplot(
      aes(x=!!sym(wage_type), y=net_migration)
      ) +
        #geom_point(alpha=.05) +
        geom_smooth(method = "lm",
                    mapping = aes(weight = !!sym(weights.)),
                    se = FALSE) +
        geom_point(data = binscatter_output$df_bin,
                   aes(x, y)) +
        coord_cartesian(ylim=c(-ylims,ylims))+
        fte_theme() +
        labs(x = x_label,
             y = "Net Migration",
             title=get_model_as_title(tidy_model, group=skill.))
}
```

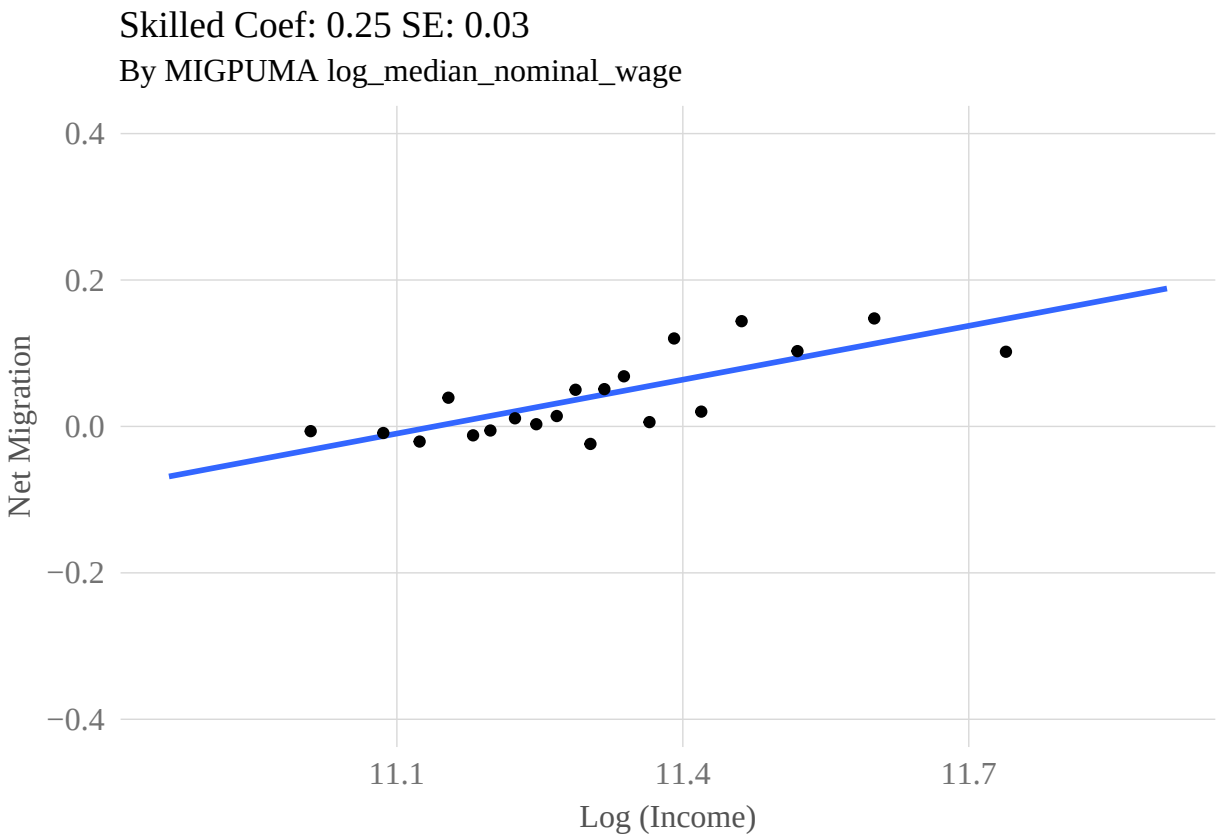
```
migpuma_high <- read_csv(file.path(working_path, "/out/high_skill_migration_by_migpuma.csv")) %>% prepa
migpuma_low <- read_csv(file.path(working_path, "/out/low_skill_migration_by_migpuma.csv")) %>% prepare_
msa_high <- read_csv(file.path(working_path, "/out/high_skill_migration_by_msa.csv")) %>% prepare_data(
msa_low <- read_csv(file.path(working_path, "/out/low_skill_migration_by_msa.csv")) %>% prepare_data("U
```

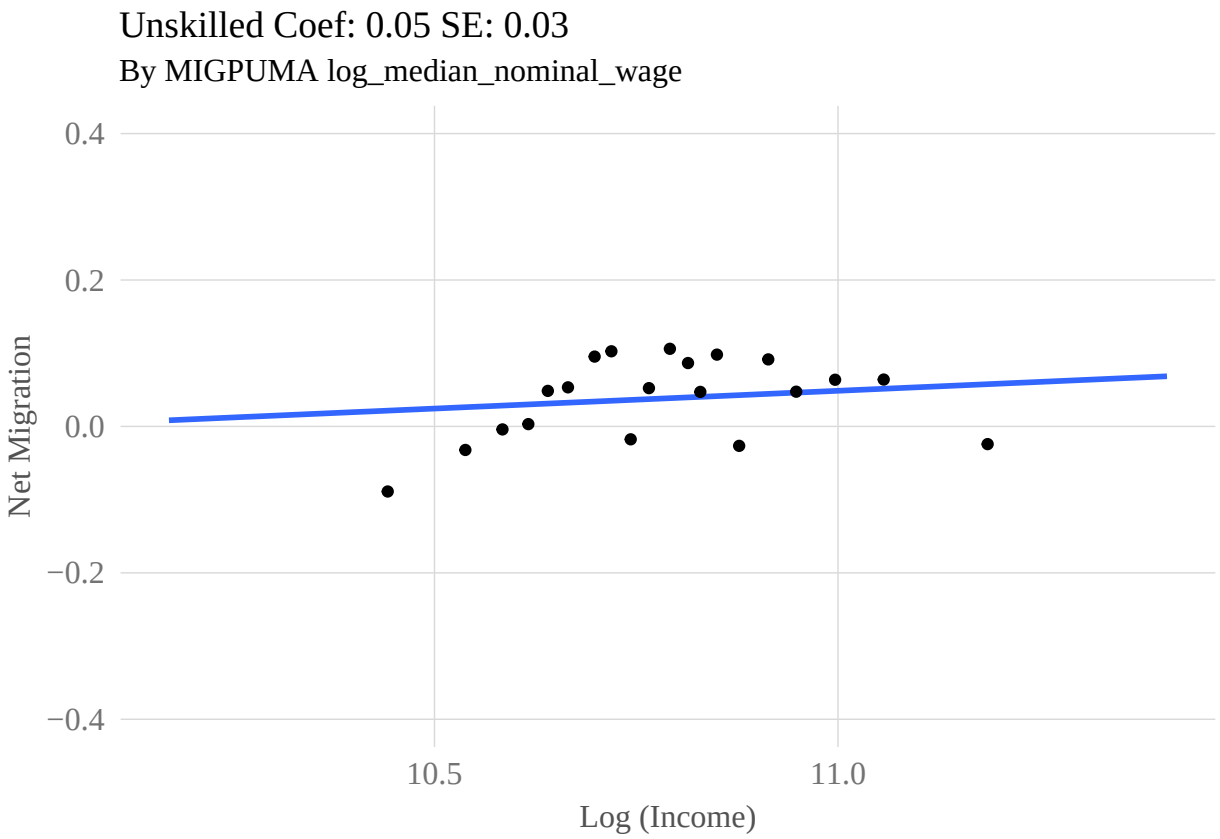
```
migpuma_plot <- bind_rows(migpuma_high, migpuma_low)
wage_list <- migpuma_plot %>% select(contains("log")) %>% names()

for (wage in wage_list){
  for (skill in c("Skilled", "Unskilled")) {
    print(migpuma_plot %>% make_plot(skill. = skill, wage_type= wage, data = .) +
            labs(subtitle = str_c("By MIGPUMA ", wage)))
  }
}
```

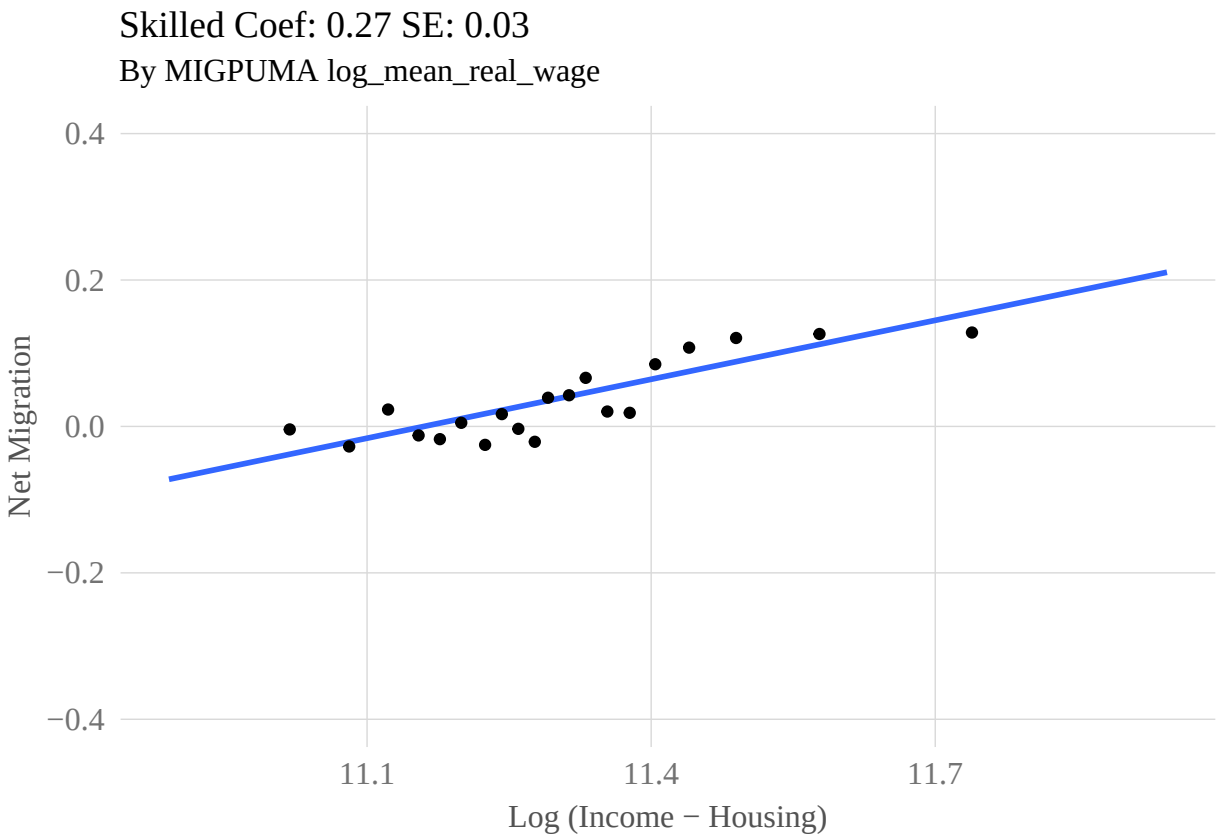
```
## Loading required package: multiwayvcov
```

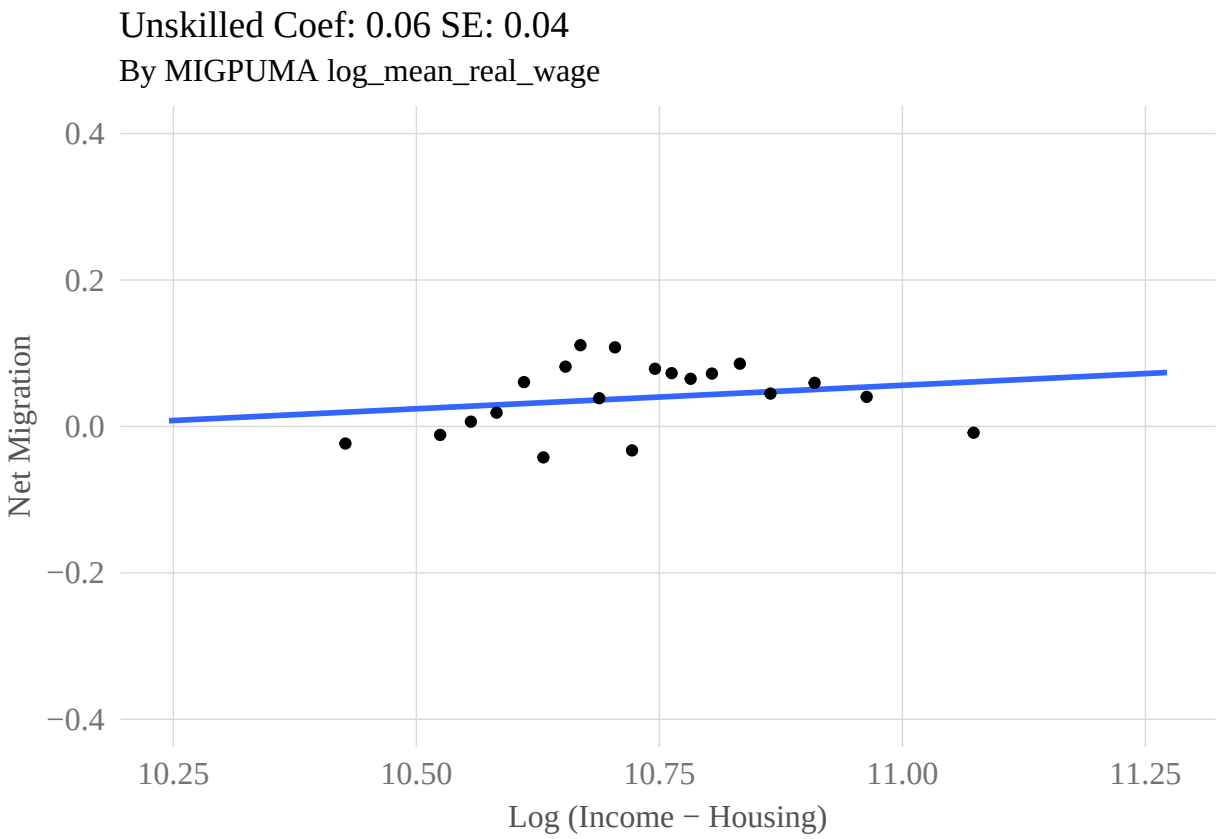


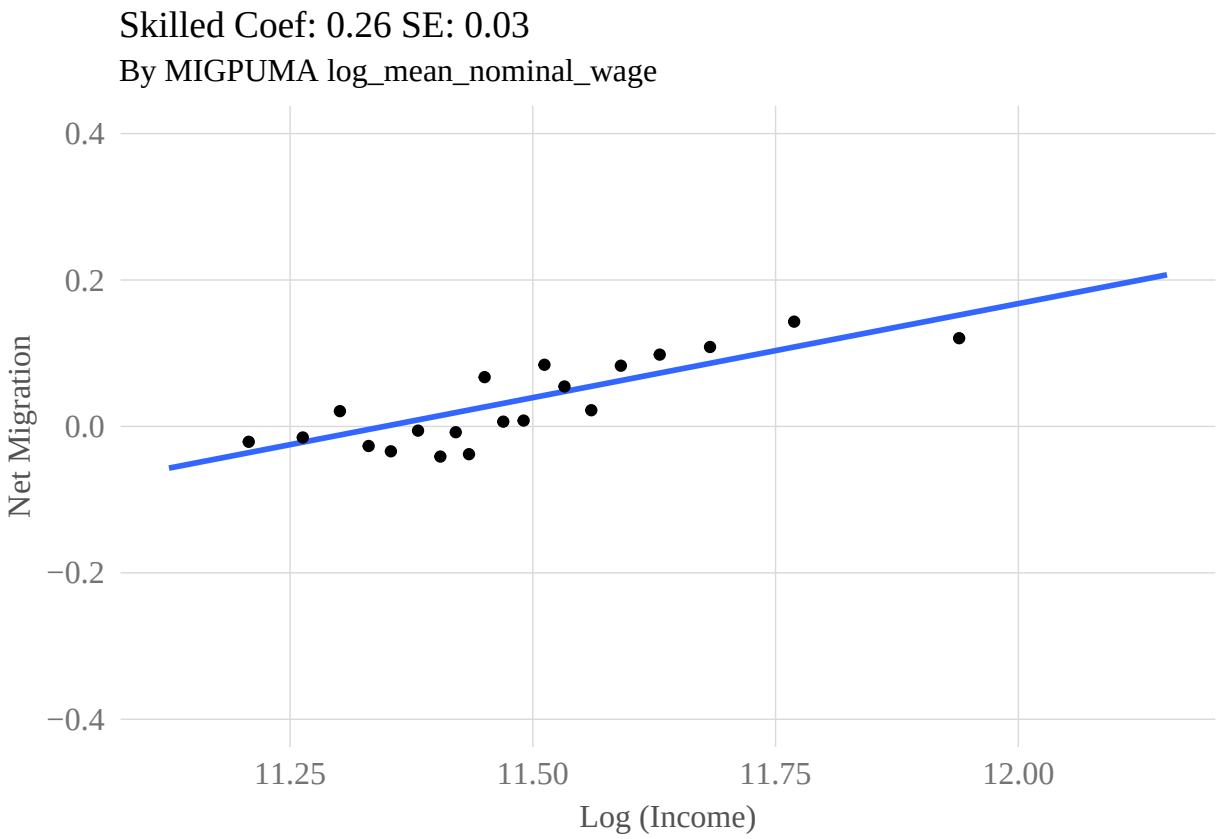

Unskilled Coef: 0.1 SE: 0.03
By MIGPUMA log_median_real_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
10.0
10.4
10.8
Log (Income − Housing)

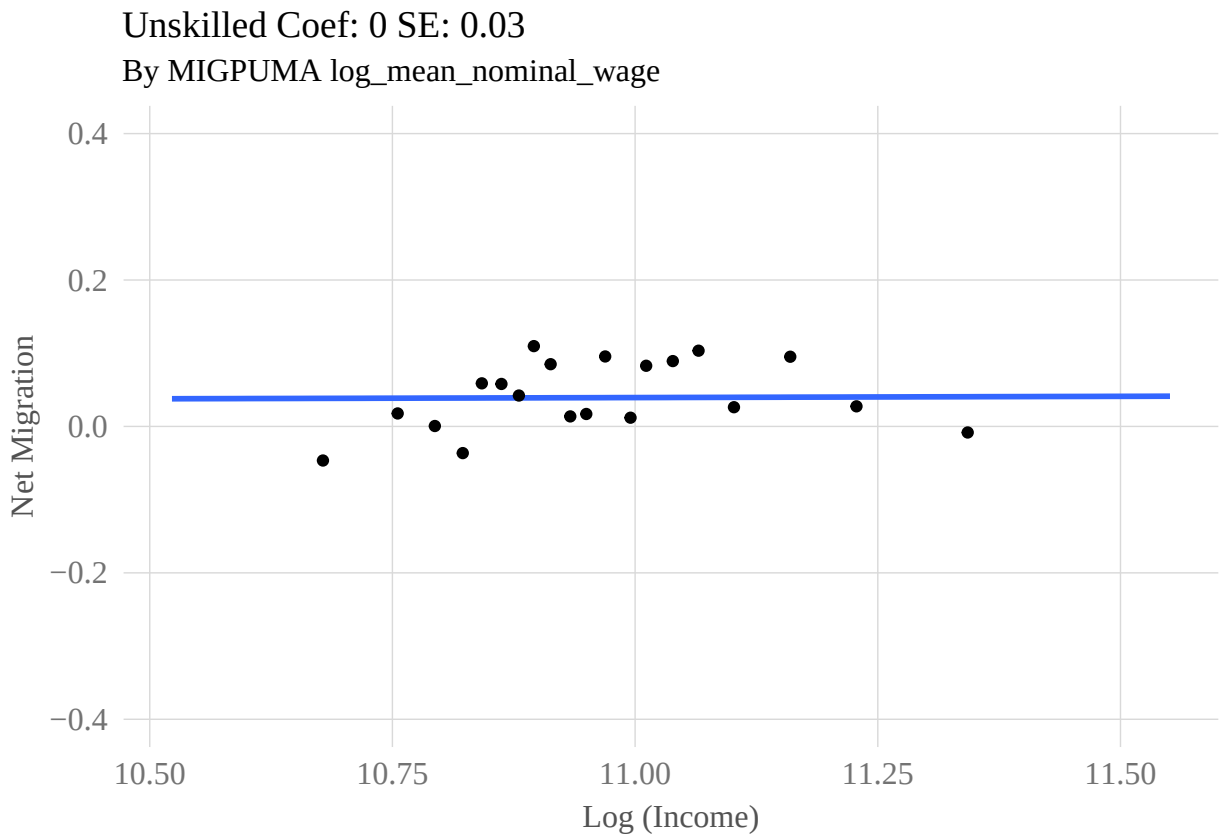

Skilled Coef: 0.25 SE: 0.03
By MIGPUMA log_median_nominal_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
11.1
11.4
11.7
Log (Income)


Unskilled Coef: 0.05 SE: 0.03
By MIGPUMA log_median_nominal_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
10.5
11.0
Log (Income)


Skilled Coef: 0.27 SE: 0.03
By MIGPUMA log_mean_real_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
11.1
11.4
11.7
Log (Income − Housing)

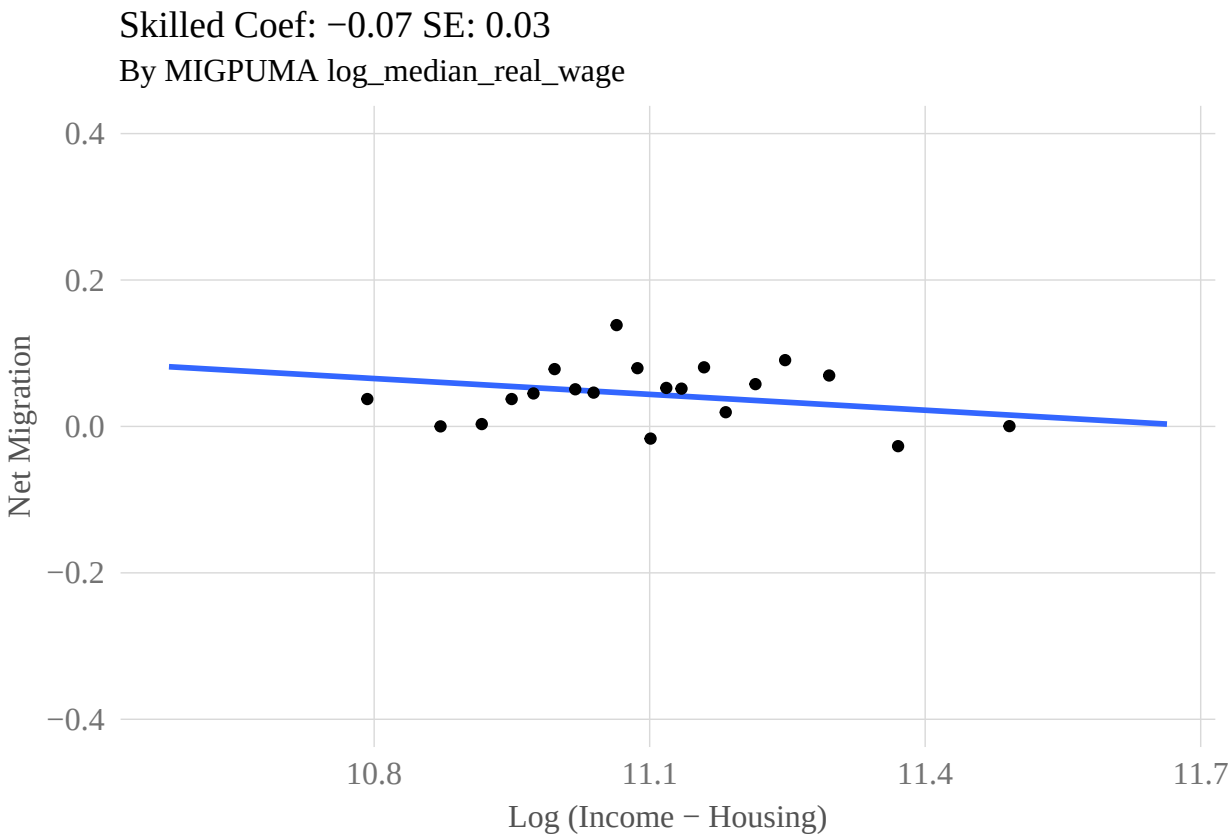

Unskilled Coef: 0.06 SE: 0.04
By MIGPUMA log_mean_real_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
10.25
10.50
10.75
11.00
11.25
Log (Income − Housing)

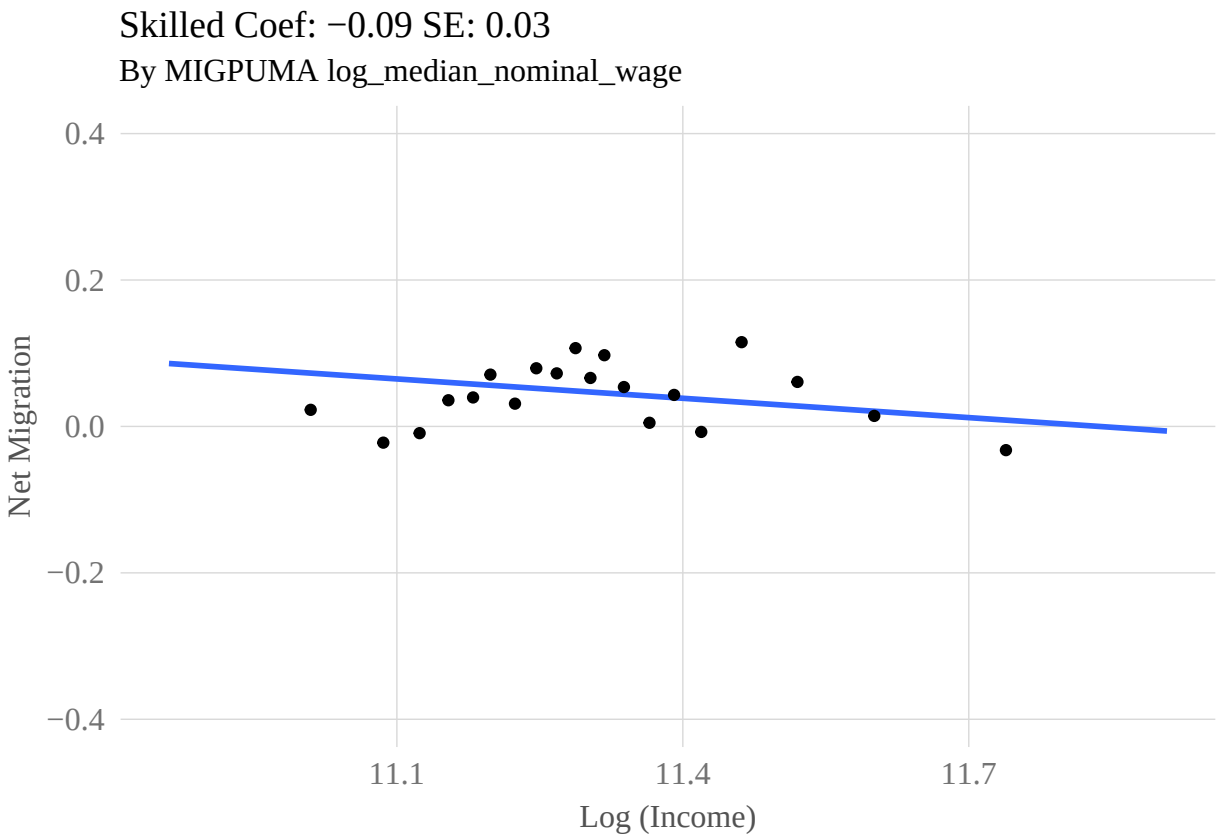

Skilled Coef: 0.26 SE: 0.03
By MIGPUMA log_mean_nominal_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
11.25
11.50
11.75
12.00
Log (Income)

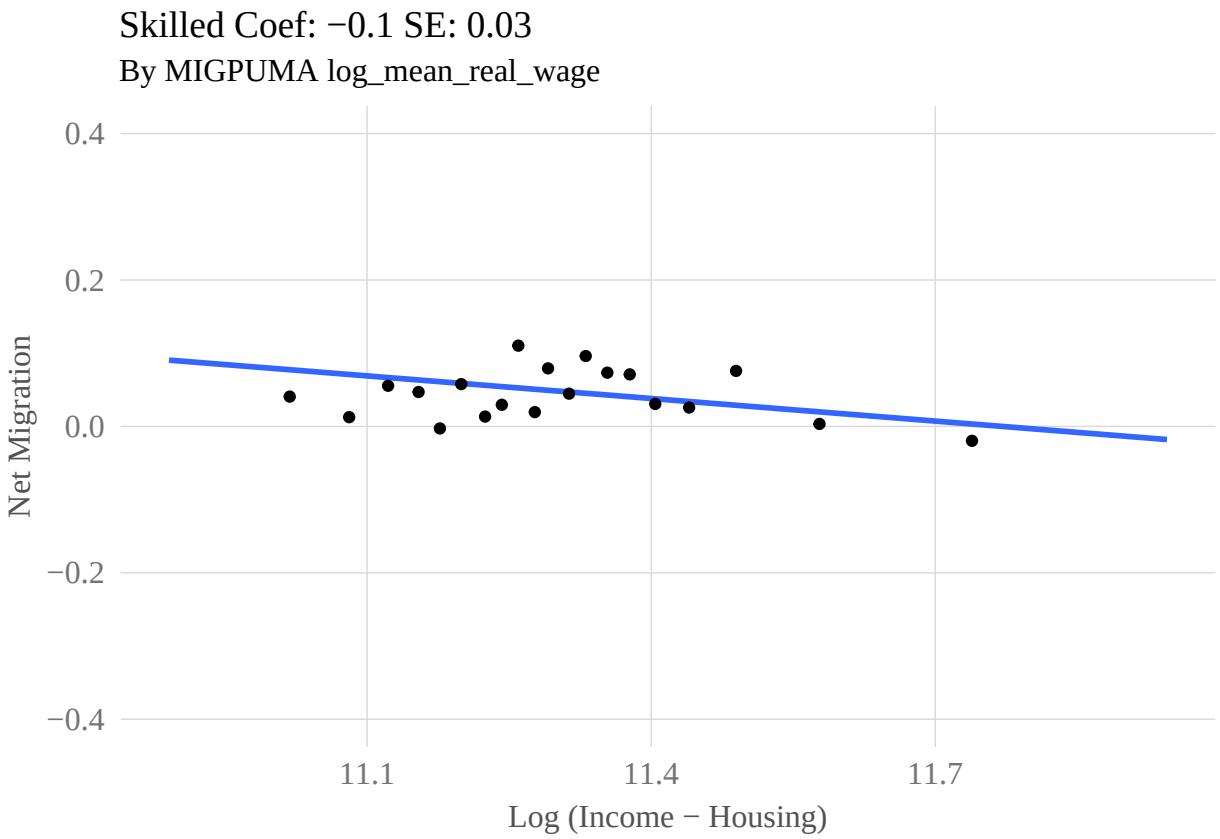


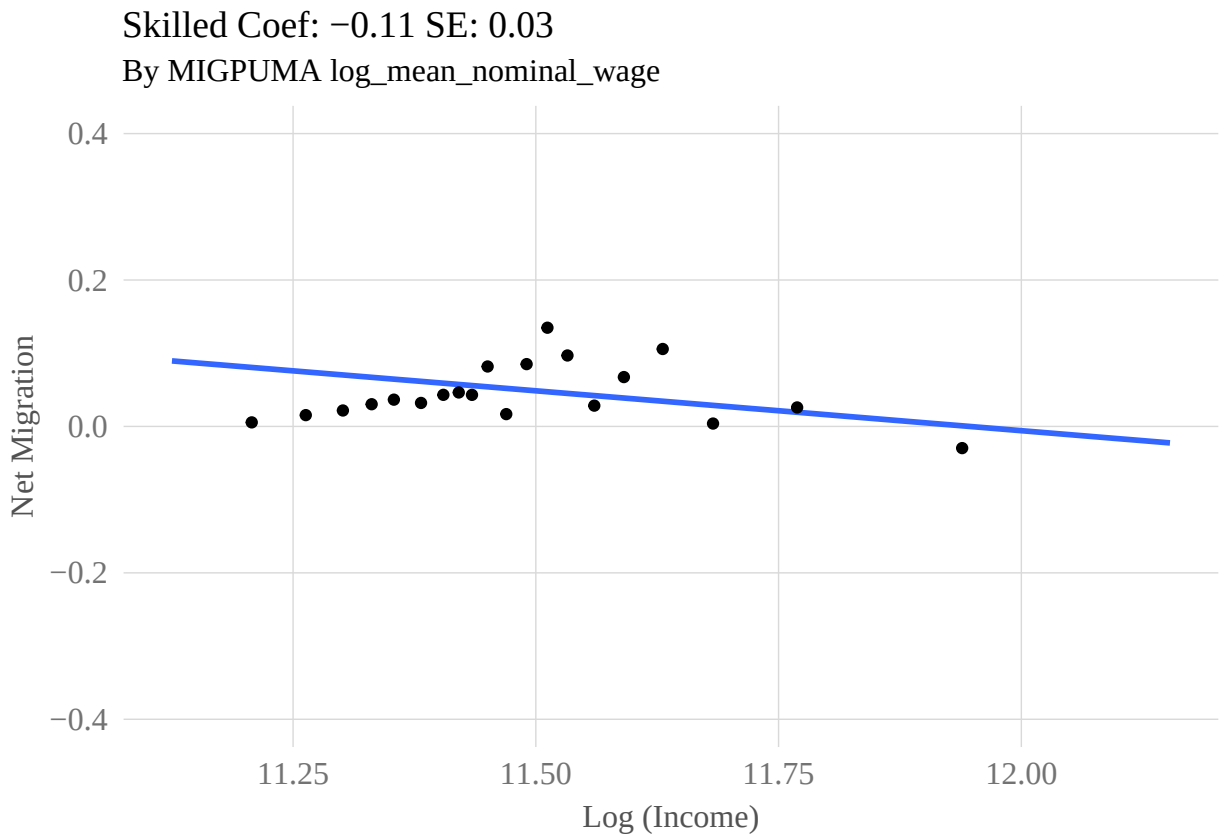
Are low-skilled workers moving to places with high high-skilled wages?

```
# y-variable: low-skill migration
# x-variable: real wages for high-skill workers

joint_migpuma_plot <-
  migpuma_high %>%
    select(state, migpuma1, migpuma_population,
           starts_with("log")) %>%
    mutate(skill = "Skilled") %>%
  left_join(
    migpuma_low %>%
      select(state, migpuma1, net_migration),
    by = c("state", "migpuma1")
    )

for (wage in wage_list){

    print(joint_migpuma_plot %>% make_plot(wage_type= wage, data = .) +
     labs(subtitle = str_c("By MIGPUMA ", wage)))
}
```


Skilled Coef: −0.07 SE: 0.03
By MIGPUMA log_median_real_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
10.8
11.1
11.4
11.7
Log (Income − Housing)


Skilled Coef: −0.09 SE: 0.03
By MIGPUMA log_median_nominal_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
11.1
11.4
11.7
Log (Income)


Skilled Coef: −0.1 SE: 0.03
By MIGPUMA log_mean_real_wage
0.4
0.2
0.0
−0.2
−0.4
Net Migration
11.1
11.4
11.7
Log (Income − Housing)



Alternative for MSAs

```
# msa_plot <-
#   msa_high %>%
#     prepare_data("Skilled", "MSA")%>%
#   bind_rows(
#     msa_low %>%
#       prepare_data("Unskilled", "MSA")
#   )
#
# # have not accounted for MSAs that were not assigned MIGPUMAs.
# msa_plot %>% make_plot(skill. = "Unskilled", data = .,  weights. = "MSA 2010 Population") + labs(subt
# msa_plot %>% make_plot(skill. = "Skilled", data = .,  weights. = "MSA 2010 Population") + labs(subtit
```