Data Scoping: Job Search Behaviour

Ebba Mark

2025-06-30

Overview

The following document summarises current progress on identifying data sources to inform the job search behaviour in our labour market ABM.

Goals:

1. Identify parameters relevant to agent search behaviour in the ABM.
2. Assess data quality for deriving empirical estimates of these parameters.

We have narrowed the list of behavioural adjustments to the following:
- **Duration-dependent search effort**
- **Reservation Wage Adjustment Rates**
- **Cyclical On-the-Job Search**
- **Risk Aversion**: For now this is randomised to ensure variation in vacancy targeting by similar workers. This is not yet supported by data.

Relevant Analyses

Data we have decided to keep:

1. Current Population Survey 2018 & 2022: Information on applications sent by unemployment duration.
2. Mukoyama et al. data on the intensive margin of unemployed search effort (in minutes searched) over the business cycle. We have chosen to include this as a validation exercise of our application effort imposition.
3. Eeckhout et al. 2019 Unemployment Cycles: We derive the sensitivity of employed job seekers to the business cycle from the employment-to-employment transitions data as used in Eeckhout et al. Due to unreliable component parts of the Eeckhout analysis, we decided to abandon using their estimated parameters (search intensity for employed workers).
4. Displaced Worker Supplement: As part of the Current Population Survey, the US Census Bureau conducts an annual Displaced Worker Supplement in which workers who have lost their job in the last three years are asked additional questions about their unemployment experiences and (if re-employed) their re-employment conditions. From this we draw a reservation wage adjustment rate as a function of unemployment duration.
5. **Mueller et al. 2021: Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias**

Additional analyses that we have decided to exclude as data inputs due to lack of relevance or poor data quality are in the "Discarded Analyses" tab.

Applications Sent

Current Population Survey - 2018 & 2020 Supplement This 2020 “Beyond the Numbers” issue distills insights from a 2018 Supplement to the Current Population Survey. The below plots show the highlights relevant to our decision-making on the job search process. In nearly all cases, the results are “binned” into intervals (ie. number of people sending 81 or more applications or unemployment duration of between 5 and 14 weeks) which means that any line plots (or linear interpretation of the bar graph) should be done carefully. Preliminary results using the raw data are found in the next section.

- Figure 1: Shows the proportion of all individuals sending x amount of applications receiving y amount of interviews. The plot indicates a “consistent” return to sending more applications, although as demonstrated in Figure 3, the number of interviews received does not necessarily equate to receiving a job offer.
- Figure 2: Demonstrates the number of applications sent (red), interviews received (green), average interview:applicaiton ratio (blue), and probability of receiving a job offer (purple) by individuals in each category of unemployment duration. There is some indication (although, again, interpretation is difficult without the raw data) that both effort and success seems to increase and then decline with time spent in unemployment, apart from success as measured by receiving a job offer which seems to consistently decline with time spent in unemployment.
- Figure 3: Percentage of jobseekers receiving an offer seems to increase as a function of the number of applications sent, until a certain point.

```
## Processing URL: https://www.bls.gov/opub/btn/volume-9/how-do-jobseekers-search-for-jobs.htm#_edn2
```

It turns out that the 2018 supplement was also run in 2022, giving us two sets of years to compare (including pre- and post-Covid). The below looks at the raw data that underlies the plotting immediately above, plus the additional data from 2022. Below find a preliminary scatter plot of applications sent versus unemployment duration. Each individual is asked how many applications they sent in the last two months (two-month periods are indicated by the grey gridlines, for reference). This does NOT include data in on the job search.

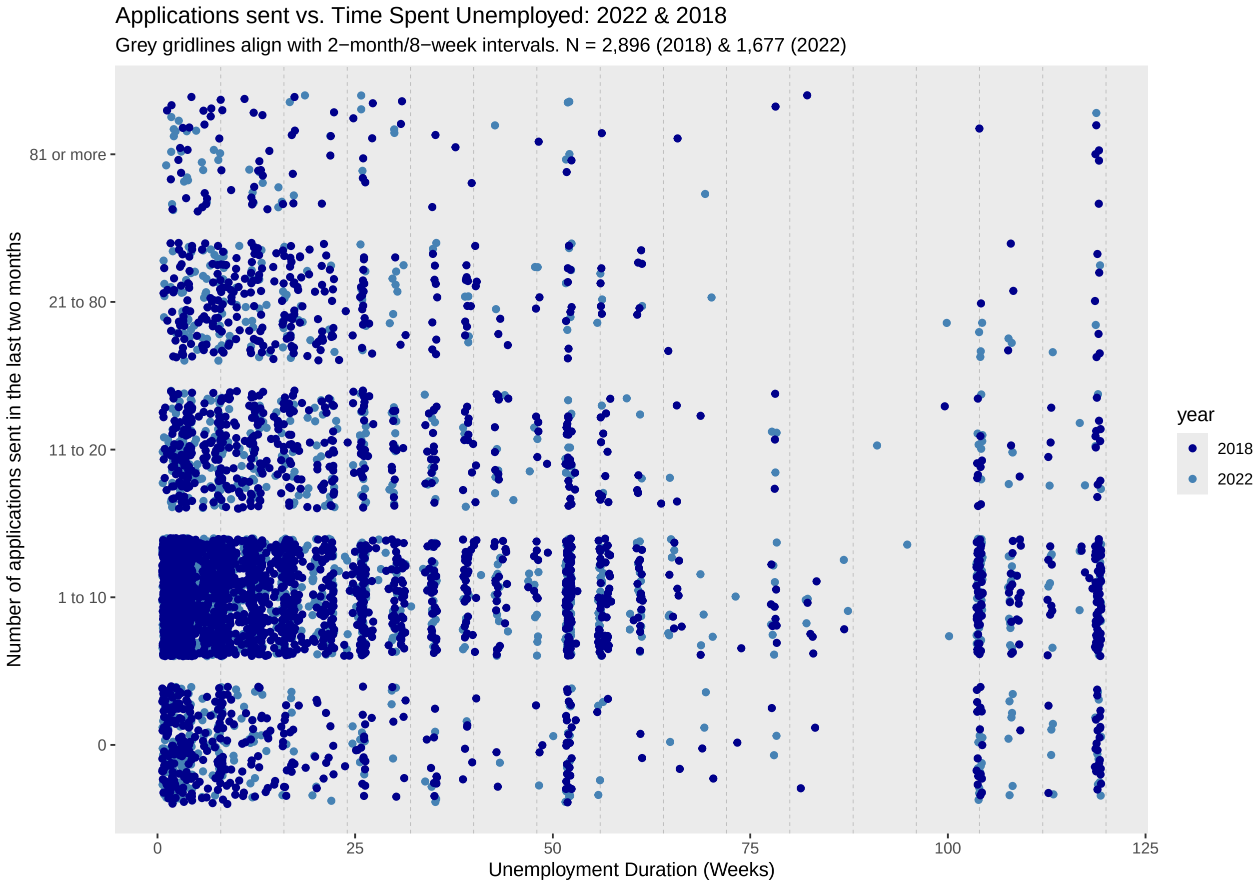
Data Source: Unemployment Insurance Nonfilers Supplement conducted in 2018 (n = 3,268) & 2022 (n = 1,901) where individuals who are unemployed but have not filed for unemployment insurance are asked the following:

PRUNEDUR	3	DURATION OF UNEMPLOYMENT FOR LAYOFF AND LOOKING RECORDS EDITED UNIVERSE: PEMLR = 3-4 <u>VALID ENTRIES</u> 0 MIN VALUE 119 MAX VALUE

Figure 1: Survey Question: Unemployment Duration

A1	Now, we also have a few questions about your experience looking for a new job over the last 2 months. How many jobs (have you/has name) applied for, if any, in the last 2 months? (Do not read the answer choices aloud) (0) 1 (1) 1 to 10 (2) 11 to 20 (3) 21 to 80 (4) 81 or more

Figure 2: Survey Question: Applications Sent



Below, I display the results of an exploration of the probability of reporting a specific number of applications sent (in the bins as in the survey question above) using various specifications of an ordinal logistic regression. I test specifications varying three different model parameters: 1. link function
2. linear vs. quadratic unemploymentduration,
3. with and without demographic control variables (education, gender, age, family income - race excluded because of lack of statistical significance though this can be revisited.)

We estimate an ordinal logistic regression model for reported applications sent $Y_i in 0, 1, 2, 3, 4$ testing four different link functions: the complementary log-log (cloglog), logistic, log-log, and probit link functions. Let $X_i^\top \beta$ denote the predictor variable. The cumulative probability of observing response category j or below, $\Pr(Y_i \leq j \mid X_i)$, is modeled as follows for each link function:

$$\begin{aligned}
\text{Complementary log-log (cloglog):} \quad & \Pr(Y_i \leq j \mid X_i) = 1 - \exp\left(-\exp\left(\tau_j - X_i^\top \beta\right)\right) \\
\text{Logistic (logit):} \quad & \Pr(Y_i \leq j \mid X_i) = \frac{1}{1 + \exp\left(-(\tau_j - X_i^\top \beta)\right)} \\
\text{Loglog:} \quad & \Pr(Y_i \leq j \mid X_i) = \exp\left(-\exp\left(-(\tau_j - X_i^\top \beta)\right)\right) \\
\text{Probit:} \quad & \Pr(Y_i \leq j \mid X_i) = \Phi(\tau_j - X_i^\top \beta)
\end{aligned}$$

Here, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The estimated coefficients β are interpreted conditional on the choice of link function where X_i is either:

$X_i = \left(\text{Unemp.Dur.}_i\right)$

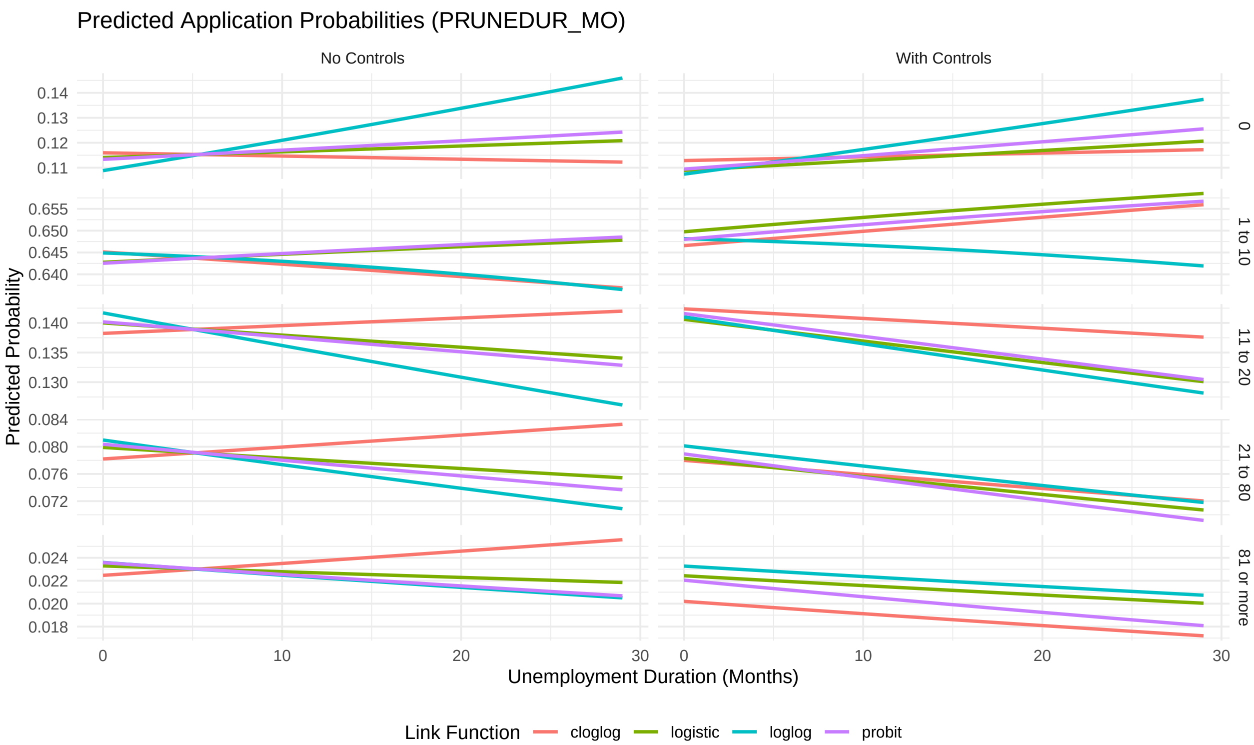
$X_i = \left(\text{Unemp.Dur.}_i^2\right)$

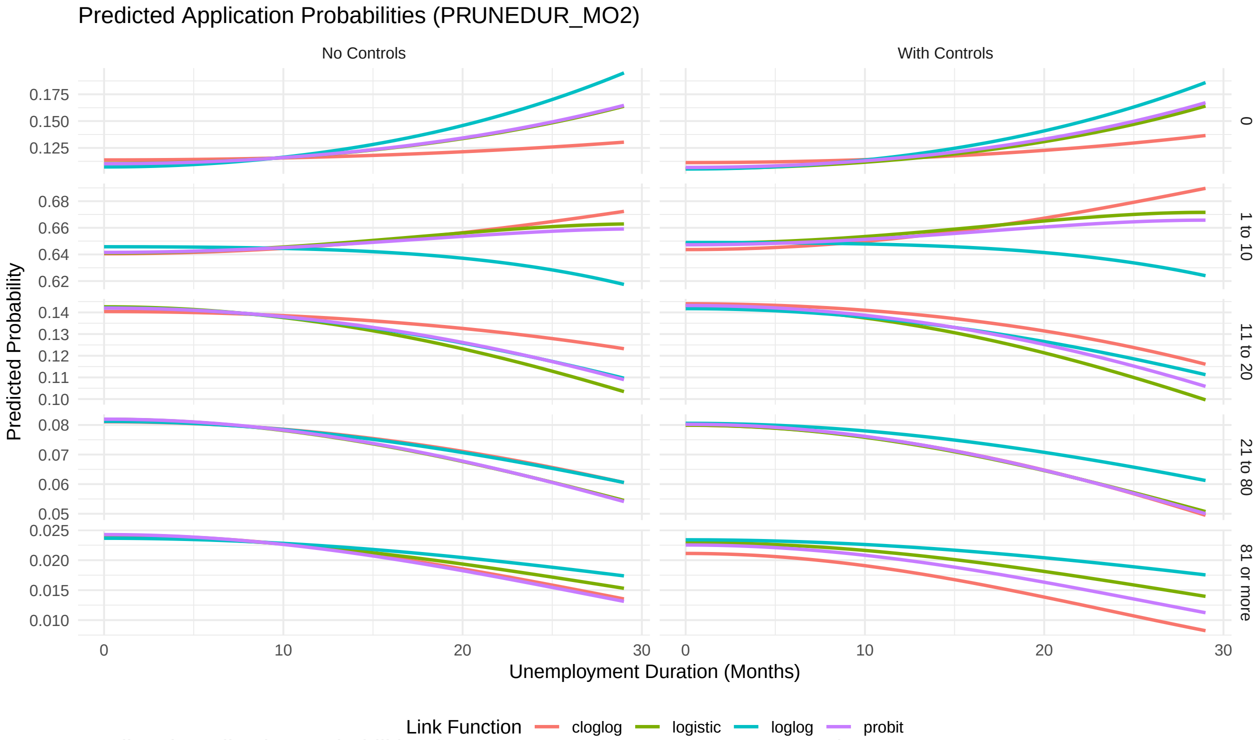
$X_i = \left(\text{Unemp.Dur.}_i, \text{Unemp.Dur.}_i^2\right)$

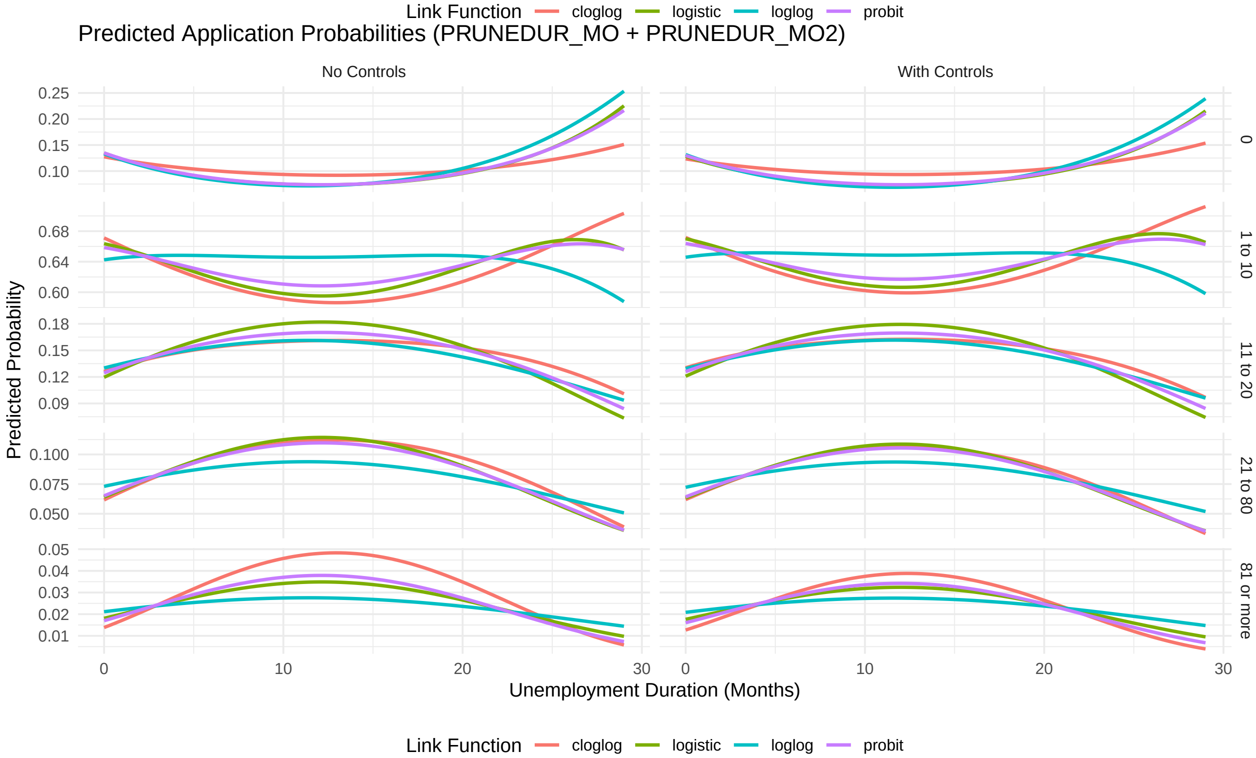
with and without control variables (education, gender, age, family income).

Assumptions about the probability distribution of the errors associated with each link function:
- *Logit:* good when the response behavior is symmetric around the middle category.
- *Probit:* When you're assuming a normal latent error distribution or want closer fit to Gaussian processes.
- *Complementary log-log:* When the likelihood of being in a higher category increases sharply but asymmetrically, or you expect hazard-like dynamics.
- *Log-log:* When early categories are of more importance and need sharper separation.

Preliminary hypothesis: Best fit will be with a complementary log-log as we care more about distinguishing between lower-level bins and there are few observations in the highest-level bins.



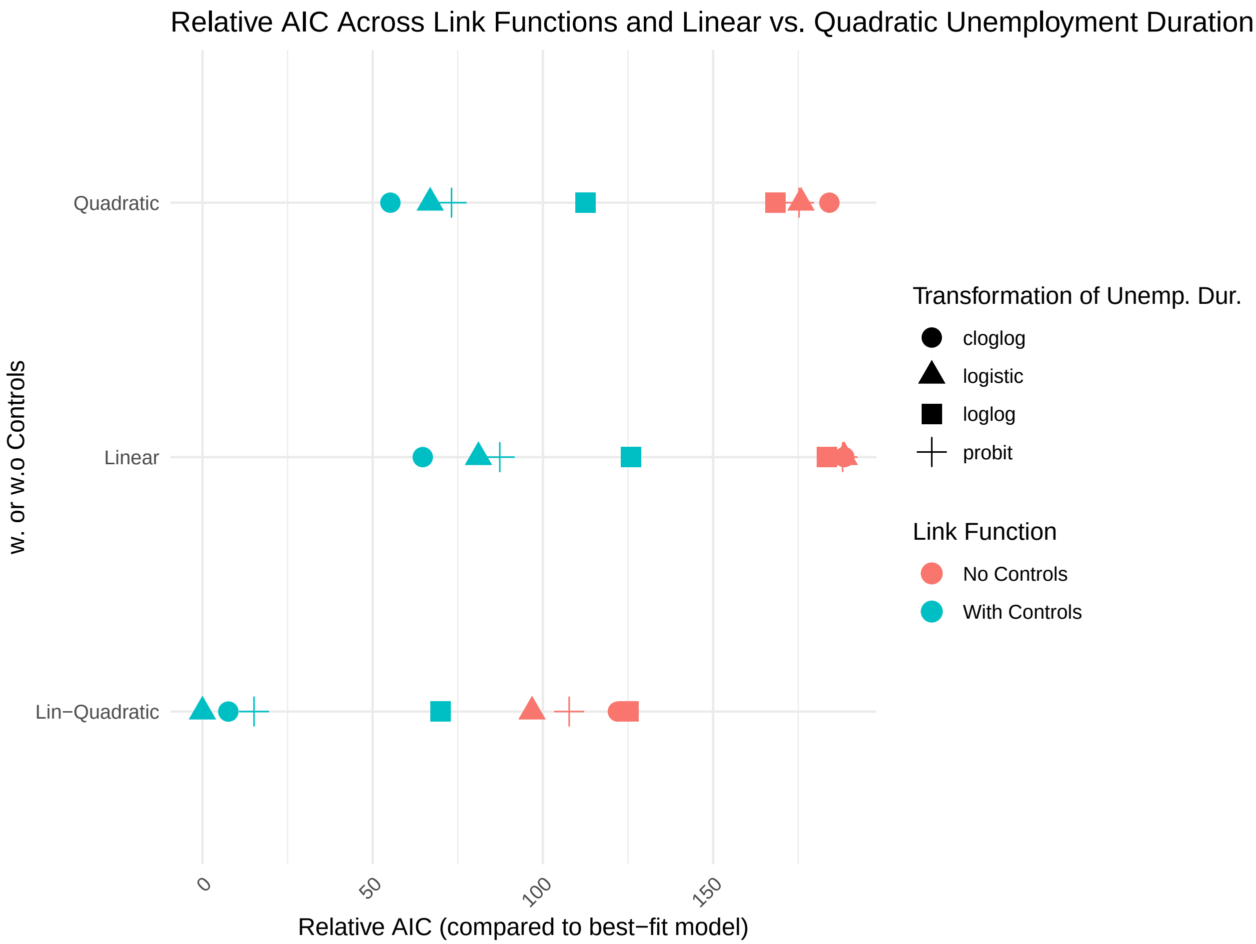




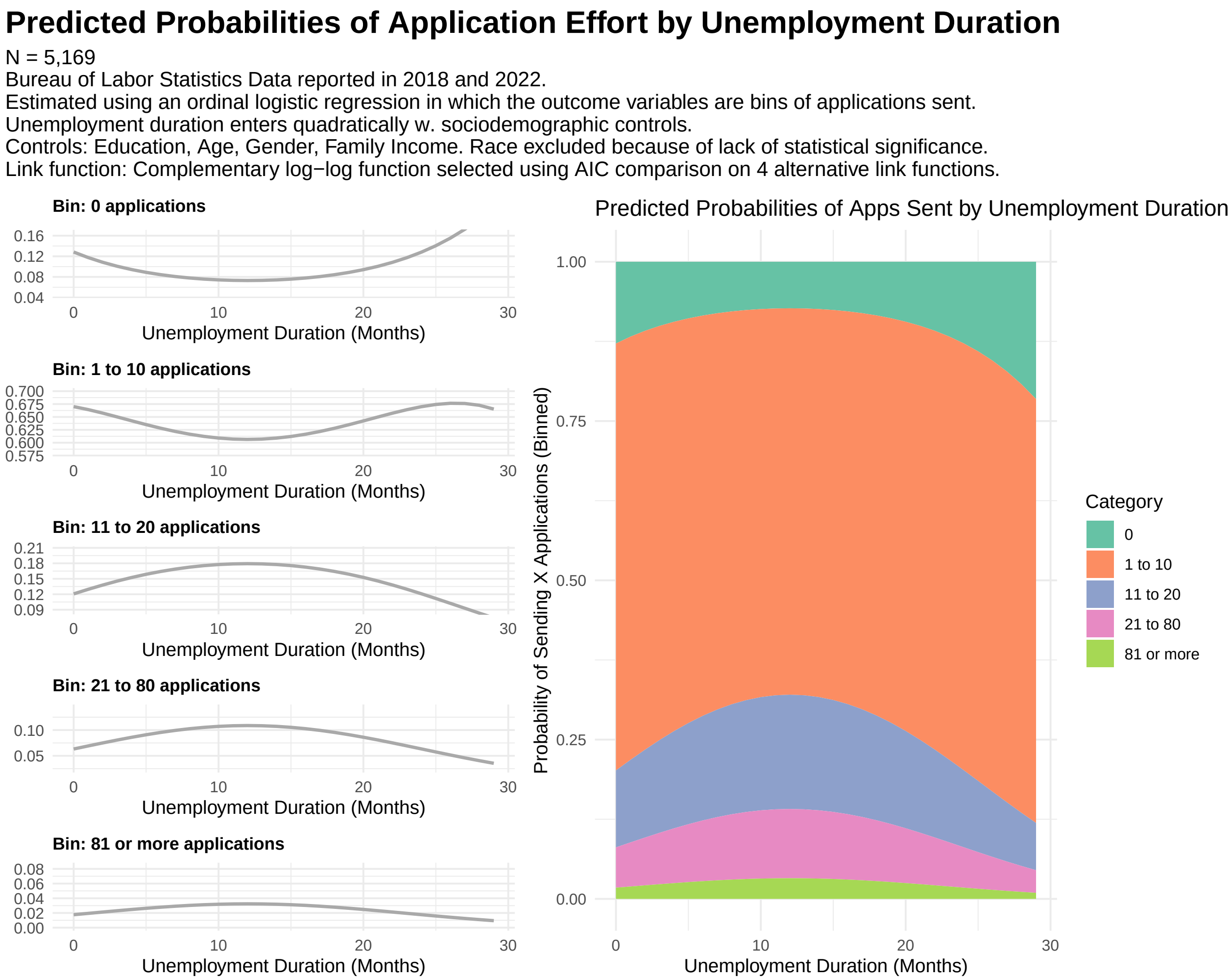
Using an AIC information criterion to compare the fit across all models, the following results are clear:
1. Models with control variables consistently perform better than those without.
2. Looking at the plots above, the relationship between unemployment duration and the predicted probability of reporting each application effort bin is very consistent except in the case of the log-log link function (blue in the panels above). In the plot below comparing the AIC the log-log link function (represented by the square symbol below) is consistently worse than all other link functions. This indicates consistency in the results reported above. Intuitively, the log-log link function is likely to be an unreasonable fit for the

latent variable as we care more about shifts in the lower-level categories than higher-level categories.
3. A complementary log-log specification for the latent variable is most suitable. This follows logically from the fact that the probability of being in the highest-level categories is relatively low.
4. Finally, a specification with a linear and quadratic estimator is consistently better than either the specification with simply a linear OR quadratic unemployment duration estimator indicating that the probability distributions represented in the final panel above are likely to be the best fit.

Result: For each additional quarter of unemployment, an individual's odds of dropping to a lower-level application category decreases by ∼.1%. This is statistically significant across all specifications at the 0.1% level.

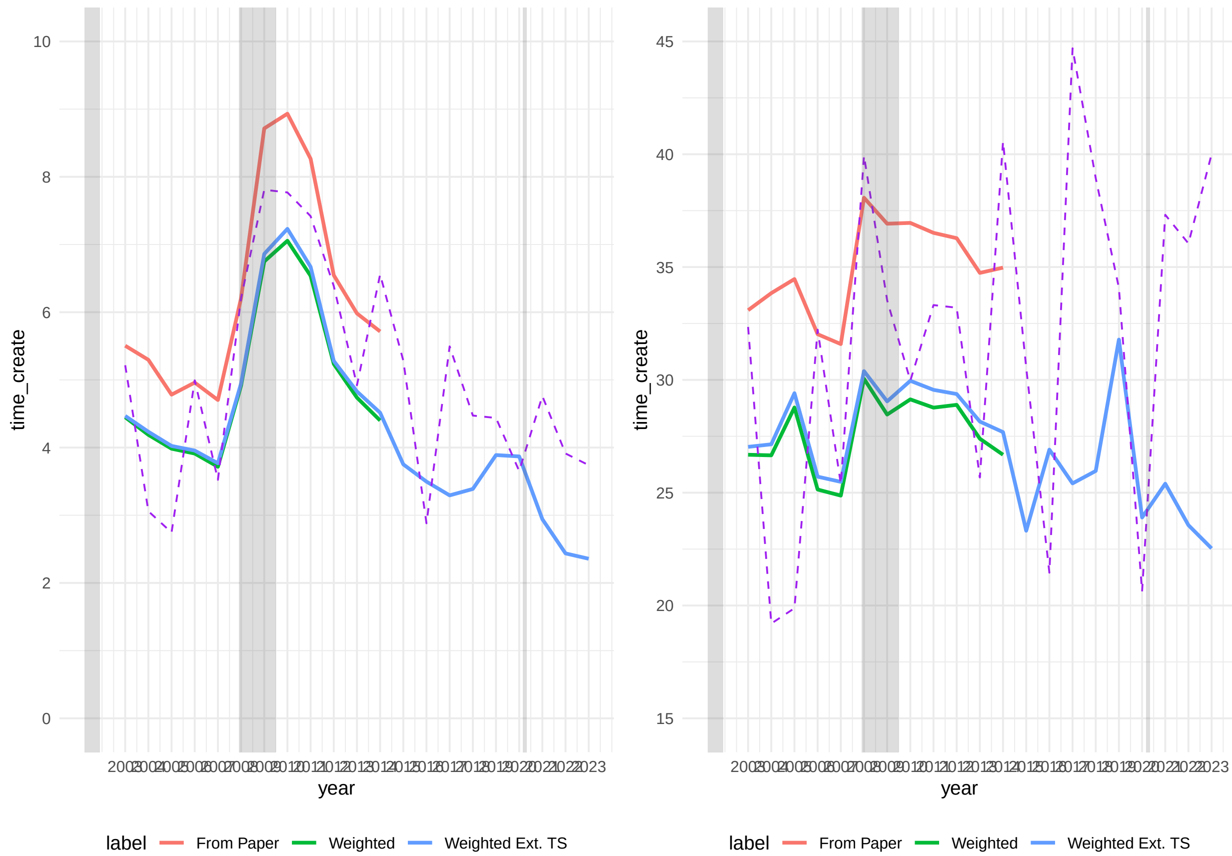



```
## [1] TRUE

## [1] TRUE
```



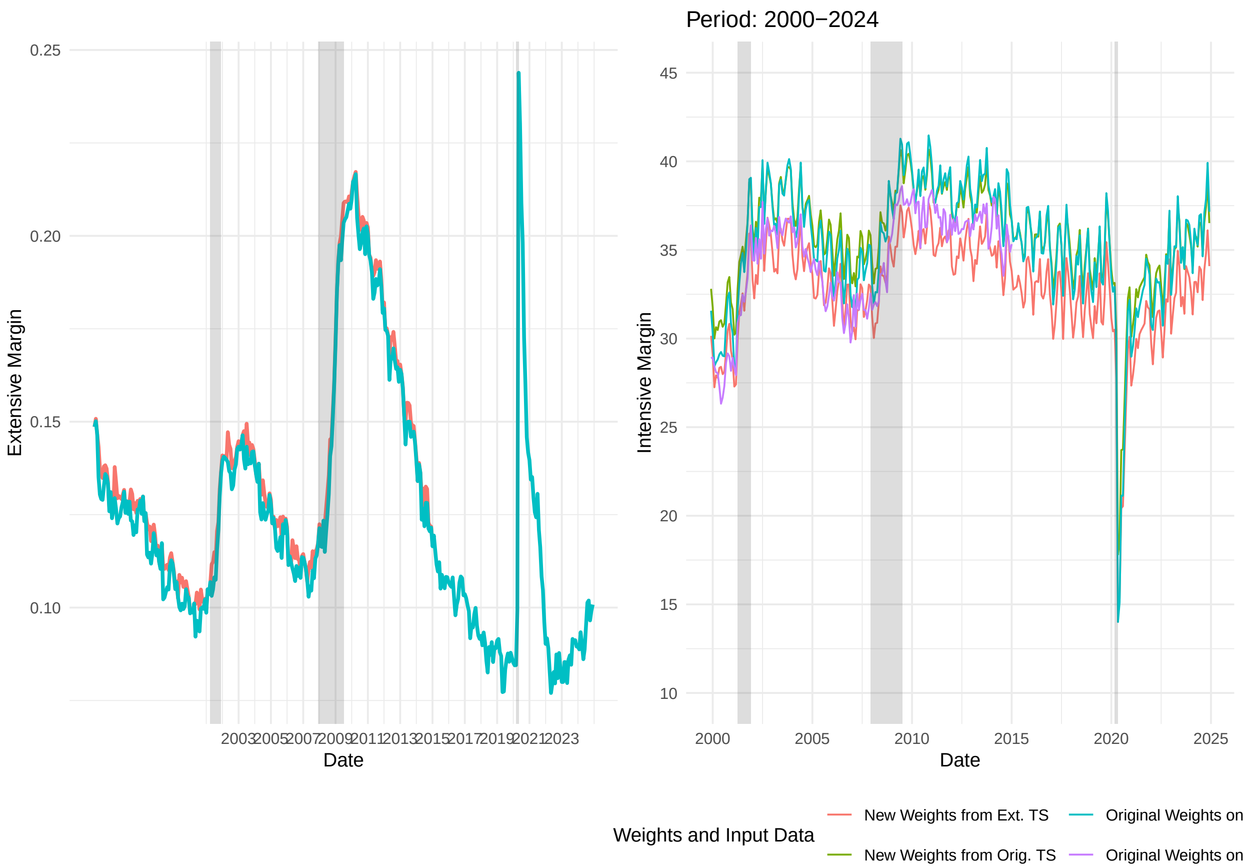
Validation: Mukoyama et al. Job Search and the Business Cycle Mukoyama: Job Search Over the Business Cycle

Figure 2. Actual and Imputed Average Search Time (minutes per day)
for All Nonemployed Workers (panel A) and Unemployed Workers (panel B)



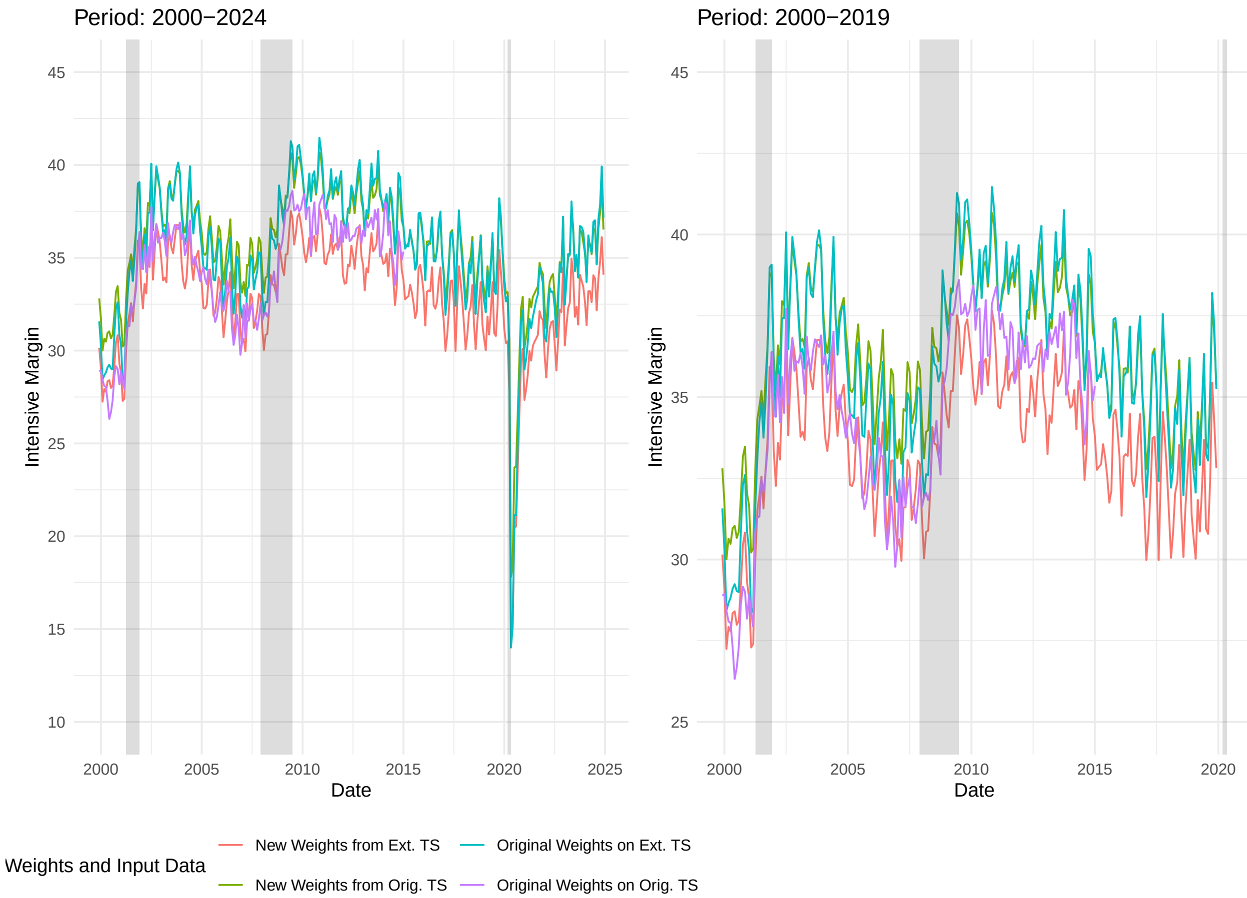
Notes: Regressions are estimated in the ATUS from 2003−2014.
A and B plot the fitted values from the sample regression, panel A plots the actual and imputed search time for all nonemployed, while panel B plots them for just the unemployed.
Observations are weighted by their ATUS sample weight.

Figure 3. The Time Series of the Extensive Margin (U/(U + N)) (panel A)
and the Intensive Margin (panel B),
Measured by the Average Minutes of Search per Day for Unemployed Workers



Red data is new data. Notes: Panel A plots the monthly ratio of the number of unemployed (U) to the total number of unemployed (U + N) in the CPS from 1994−2014.

Intensive Margin Measured by the Average Minutes of Search per Day for Unemployed Workers



Plots the average minutes of search per day,using the imputed minutes as a function of search methods used.
Each observation is weighted by its CPS sample weight.

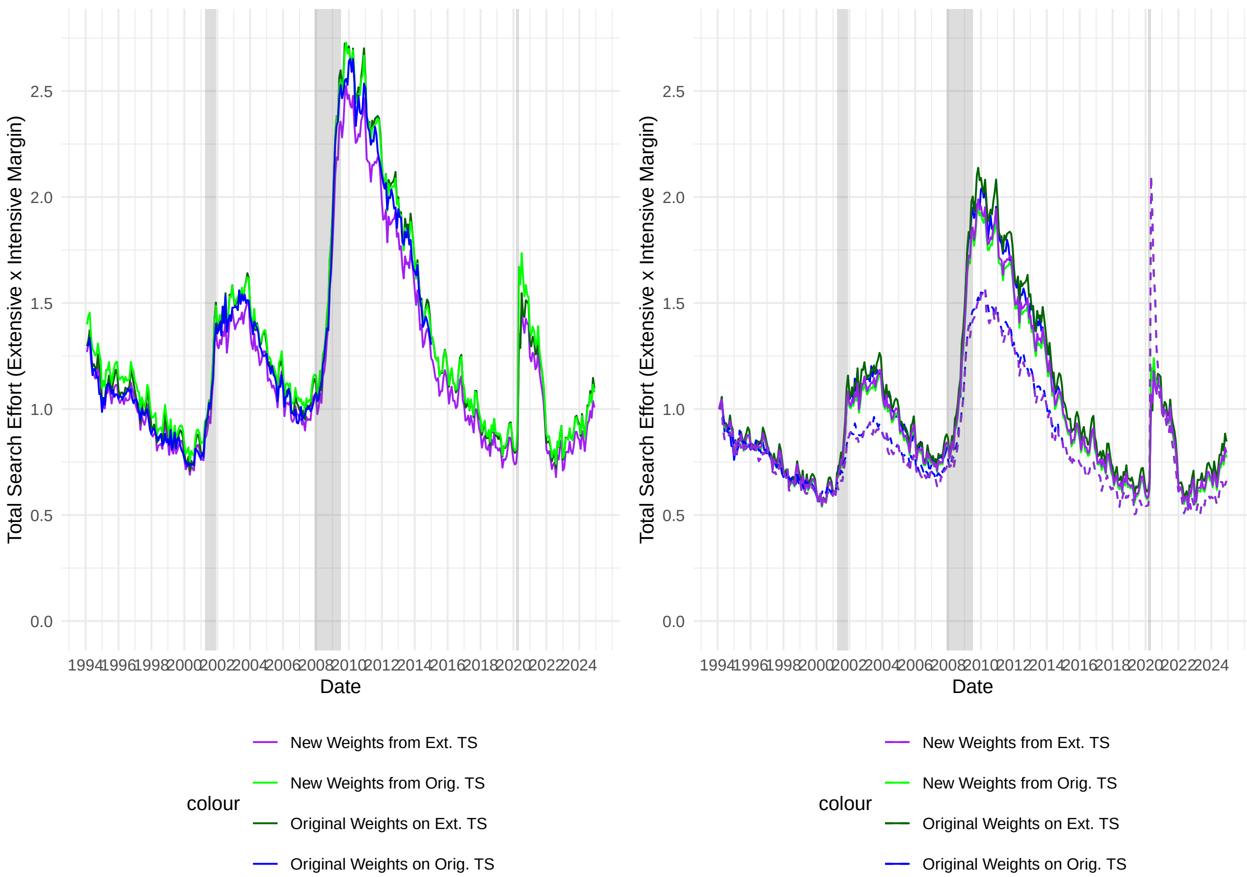


Figure 4. Time Series of (Panel A) Total Search Effort and (Panel B) Total Search Effort Using the Search Time of Unemployed Workers [solid: (s*(U/(E + U + N))] versus Using the Number of Unemployed Workers [dashed: U/(E + U + N)) (panel B)

Reservation Wage Adjustment

As part of the Current Population Survey, the US Census Bureau conducts an annual Displaced Worker Supplement in which workers who have lost their job in the last three years are asked additional questions about their unemployment experiences and (if re-employed) their re-employment conditions.

From link above: "The universe for the Displaced Workers Supplement is civilians 20 or older. Respondents are further categorized as a"displaced worker" if they meet additional characteristics (see DWSTAT). After 1998, displaced workers are those who lost or left a job due to layoffs or shutdowns within the past 3 years…were not self-employed, and did not expect to be recalled to work within the next six months.

The data used below is from annual survey responses between 2000-2025. I use the supplement sample weights in all results below. I note where I have clipped the sample for outliers (wage ratio between [0.25, 2] and unemployment duration less than 96 weeks (~24 months).

Below I:

1. **Data Cleaning Procedure:** Show data cleaning just for reference (feel free to ignore)!.
2. **Descriptives:** Show some descriptives about the data itself.
3. **Regression Results on Non-Uniform Sample:** Regression results with ratio of new wage to wage at the lost job (W_h and W_w) regressed (cross-sectionally) on unemployment duration with and without various combinations of control variables (whether or not an individual received unemployment compensation, age, race, sex, marital status, education, previous wage level.) Note that the wages are

reported in hourly and weekly values but. this reporting is inconsistent across observations. In other words, though most individuals (4600/6198) report their wage in both units, 270 report only hourly and 1328 report only weekly. I have not reconciled the inconsistency so I use hourly wage ratios in majority of the below document. I could try to reconcile this.

4. Outline some considerations for further improvement of the analysis:

    1. **Reweighted Samples:** The sample is non-uniform in unemployment duration (less observations as unemployment duration increases). Try two methods of reweighting to address selection issues (Heckman Selection correction - though I think this is inappropriate for this particular selection issue) and non-uniform (entropy-balancing to deal with representativeness of population over unemployment durations) sample confirm regression results in non-uniform sample.
    2. **Representativeness of the Sample (Education, Age, Gender, and Wage):** Representativeness of the data to motivate data limitations and inform the ultimate reweighting scheme.

Overall result (at the moment): Individuals accept a ~1-percentage point change in the wage ratio per additional month of unemployment. Variations using model reweighting, different samples, combinations of control variables, reported hourly and weekly wage ratios do not seem to affect the result. However, the data seems to follow a non-linear relationship (we see little satisficing until around ~12 months of unemployment) after which the wage ratio begins to decrease. Individuals seem to accept a below-1 relative wage ratio (current wage:wage at lost job) following a year of unemployment. If we fit this model with a quadratic fit this could inform our reservation wage adjustment parameter in the model.

Important Considerations/Limitations:

1. **Displaced worker classification as outlined above.** Can we generalise from this definition to all unemployed workers?
2. **The reported 'current wage' is not necessarily the the realised wage post-re-employment.** Individuals report the wage at the lost job, the amount of time unemployed until they were re-employed, and the wage they hold at their current job. However, it is not indicated whether the current job is the same job as the first they were re-employed at. Given various comments in the literature about finding "stop-gap" employment, this might not be a problem in the sense that the "current wage" would more accurately indicate the wage an individual has "landed" at post-unemployment spell. But curious what you think about the defensibility of this.
3. **Outcome variable:** How do we feel about the outcome variable as the ratio of current to latest held job? Might we want to take the log or consider simply the (log) level regressed on the previous wage. Wondering if a ratio-based outcome variable might muddle interpretation. Curious for your reactions.

Data cleaning

Feel free to ignore this code chunk immediately below - included for now for transparency in case you spot issues. I include it for your info on binning and outlier trimming.

```r
# From the original dataset, I include only those that reported having lost a FT job in the last three
df <- readRDS(here("data/behav_params/cps_displaced_worker_supplement/cps_disp_filtered.RDS")) %>%
  select(hwtfinl, cpsid, wtfinl, age, sex, race, marst, educ, # age, sex, race, marital status, educati
         dwsuppwt, # Survey weight
         dwyears, # Years worked at lost job
         dwben, # Received unemployment benefits
         dwexben, # Exhausted unemployment benefits
         dwlastwrk, # Time since worked at last job
         dwweekc, # Weekly earnings at current job
         dwweekl, # Weekly earnings at lost job
         dwwagel, # Hourly earnings at lost job
         dwwagec, # Hourly wage at current job
```

```r
         dwhrswkc, # Hours worked each week at current job
         dwresp, # Eligibility and interview status for Displaced Worker Supplement
         # Interestingly the unemployment duration is not directly linked to CURRENT job and we cannot
         dwwksun) %>%  # Number of weeks not working between between end of lost or left job and start
  # I remove anyone who is Not in Universe (99) and declaring greater than 160 weeks unemployed between
filter(dwhrswkc != 99 & dwwksun <= 160) %>%
  # Replacing NIU values with NA values
  mutate(dwwagel = ifelse(round(dwwagel) == 100, NA, dwwagel),
         dwwagec = ifelse(round(dwwagec) == 100, NA, dwwagec),
         dwweekl = ifelse(round(dwweekl) == 10000, NA, dwweekl),
         dwweekc = ifelse(round(dwweekc) == 10000, NA, dwweekc),
         # dwwage_rec_l = ifelse(is.na(dwagel) & !is.na(dweekl) ~ dwweekl),
         # dwweekc = ifelse(round(dwweekc) == 10000, NA, dwweekc),
         # Binning educational categories
         educ_cat = factor(case_when(educ %in% c(1) ~ NA, # (NIU)
                                     educ > 1 & educ <= 71 ~ "Less than HS", # Includes "None" - Grade 12 no d
                                     educ %in% c(73, 81) ~ "HS Diploma", # Includes "High school Diploma or eq
                                     educ %in% c(91, 92) ~ "Associate's", # Include "[Associate's degree, occu
                                     educ %in% c(111) ~ "Bachelor's", # Bachelor's degree
                                     educ > 111 ~ "Postgraduate Degree" # Includes Master's, Professional Scho
                                     ), , levels = c("Less than HS", "HS Diploma", "Associate's", "Bachelor's"
         # Marital status to binary indicator
         marst = case_when(marst == 1 ~ 1, # Married with a present spouse
                           # Might consider dividing this differently
                           TRUE ~ 0), # Married with absent spouse, separated, divorced, widowed, never
         # gender to 0,1 values
         female = sex == 2,
         # race to higher-level categories w binary values
         white = race == 100,
         black = race == 200,
         mixed = race %in% c(801, 802, 803, 804, 805, 806, 810, 812, 813, 820, 830),
         aapi = race %in% c(650, 651, 652, 808, 809),
         native = race == 300
         # age is a continuous variable which seems fine for now...binning likely unnecessary
         ) %>%
        # Ratio of hourly wage of current job to lost job
  mutate(ratio_wage = dwwagec/dwwagel,
         # Ratio of weekly wage of current job to lost job
         ratio_weekly = dwweekc/dwweekl,
         # Reconciling missing reporting between weekly and hourly wage. Take either the min, max or me
         ratio_reconciled_min = case_when(is.na(ratio_wage) ~ ratio_weekly,
                                          is.na(ratio_weekly) ~ ratio_wage,
                                          TRUE ~ pmin(ratio_weekly, ratio_wage)),
         ratio_reconciled_max = case_when(is.na(ratio_wage) ~ ratio_weekly,
                                          is.na(ratio_weekly) ~ ratio_wage,
                                          TRUE ~ pmax(ratio_weekly, ratio_wage)),
         ratio_reconciled_mean = case_when(is.na(ratio_wage) ~ ratio_weekly,
                                           is.na(ratio_weekly) ~ ratio_wage,
                                           TRUE ~ rowMeans(across(c(ratio_wage, ratio_weekly)), na.rm = T
         # Create monthly unemployment duration for continuous
         dwmosun = floor(dwwksun/4),
         dwmosun2 = dwmosun^2,
         dwmosun3 = dwmosun^3,
```

```
# Unemployment duration (reported as time between lost job and start of next job)
# I bin in...
# monthly intervals (4 weeks) from 1-6 months
# quarterly intervals (12 weeks) from 7 mos-1 year
# half-year interval from 1-2.5 years
# single bin for anyone about 120 weeks
dwwksun_bin = case_when(
  # Monthly intervals (4 weeks) from 1-6 months
  dwwksun <= 4 ~ 1, #"Less than 4 weeks",
                      dwwksun > 4 & dwwksun <= 8 ~ 2,
                      dwwksun > 8 & dwwksun <= 12 ~ 3,
                      dwwksun > 12 & dwwksun <= 16 ~ 4,
                      dwwksun > 16 & dwwksun <= 20 ~ 5,
                      dwwksun > 20 & dwwksun <= 24 ~ 6,
                      # Quarterly Intervals (12 weeks) from 6+ mos - 1 year
                      dwwksun > 24 & dwwksun <= 36 ~ 7,
                      dwwksun > 36 & dwwksun <= 48 ~ 8,
                      # Half-year Intervals (24 weeks) from 1-2.5 years
                      dwwksun > 48 & dwwksun <= 72 ~ 9,
                      dwwksun > 72 & dwwksun <= 96 ~ 10,
                      dwwksun > 96 & dwwksun <= 120 ~ 11,
                      # Anyone above - recall this is capped at 160 weeks as per filter abov
                      dwwksun > 120 ~ 12),
# Bin labels
dwwksun_bin_labs = case_when(dwwksun_bin == 1 ~ "<= 1 mo.", #"Less than 4 weeks",
                      dwwksun_bin == 2 ~ "1-2 mos.",
                      dwwksun_bin == 3 ~ "2-3 mos.",
                      dwwksun_bin == 4 ~ "3-4 mos.",
                      dwwksun_bin == 5 ~ "4-5 mos.",
                      dwwksun_bin == 6 ~ "5-6 mos.",
                      # Quarterly Intervals (12 weeks) from 6+ mos - 1 year
                      dwwksun_bin == 7 ~ "6-9 mos.",
                      dwwksun_bin == 8 ~ "9-12 mos.",
                      # Half-year Intervals (24 weeks) from 1-2.5 years
                      dwwksun_bin == 9 ~ "12-18 mos.",
                      dwwksun_bin == 10 ~ "18-24 mos.",
                      dwwksun_bin == 11 ~ "24-30 mos.",
                      # Anyone above - recall this is capped at 160 weeks as per filter abov
                      dwwksun_bin == 12 ~ "30+ mos."),
log_ratio_wage = log(ratio_wage),
log_ratio_weekly = log(ratio_weekly),
# I clip the sample to an accepted wage ratio between [0.5, 2] and less than 96 weeks of unemp
clipped_sample_hwage = ratio_wage >= 0.5 & ratio_wage <= 2 & dwwksun_bin < 11,
clipped_sample_wwage = ratio_weekly >= 0.5 & ratio_weekly <= 2  & dwwksun_bin < 11,
clipped_sample_rec_min = ratio_reconciled_min >= 0.5 & ratio_reconciled_min <= 2 & dwwksun_bin
clipped_sample_rec_max = ratio_reconciled_max >= 0.5 & ratio_reconciled_max <= 2 & dwwksun_bin
clipped_sample_rec_mean = ratio_reconciled_mean >= 0.5 & ratio_reconciled_mean <= 2 & dwwksun_
```

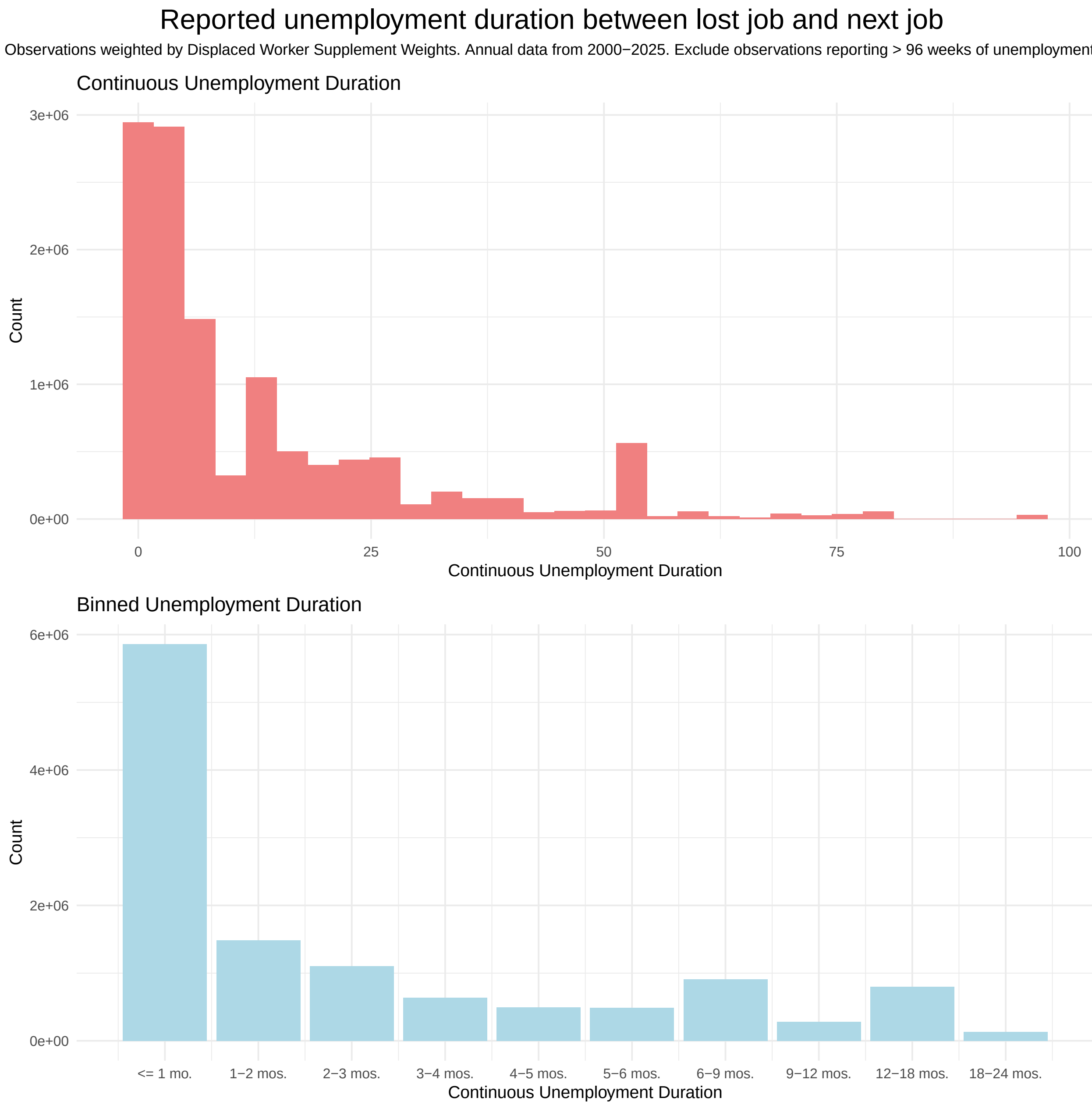
Descriptives

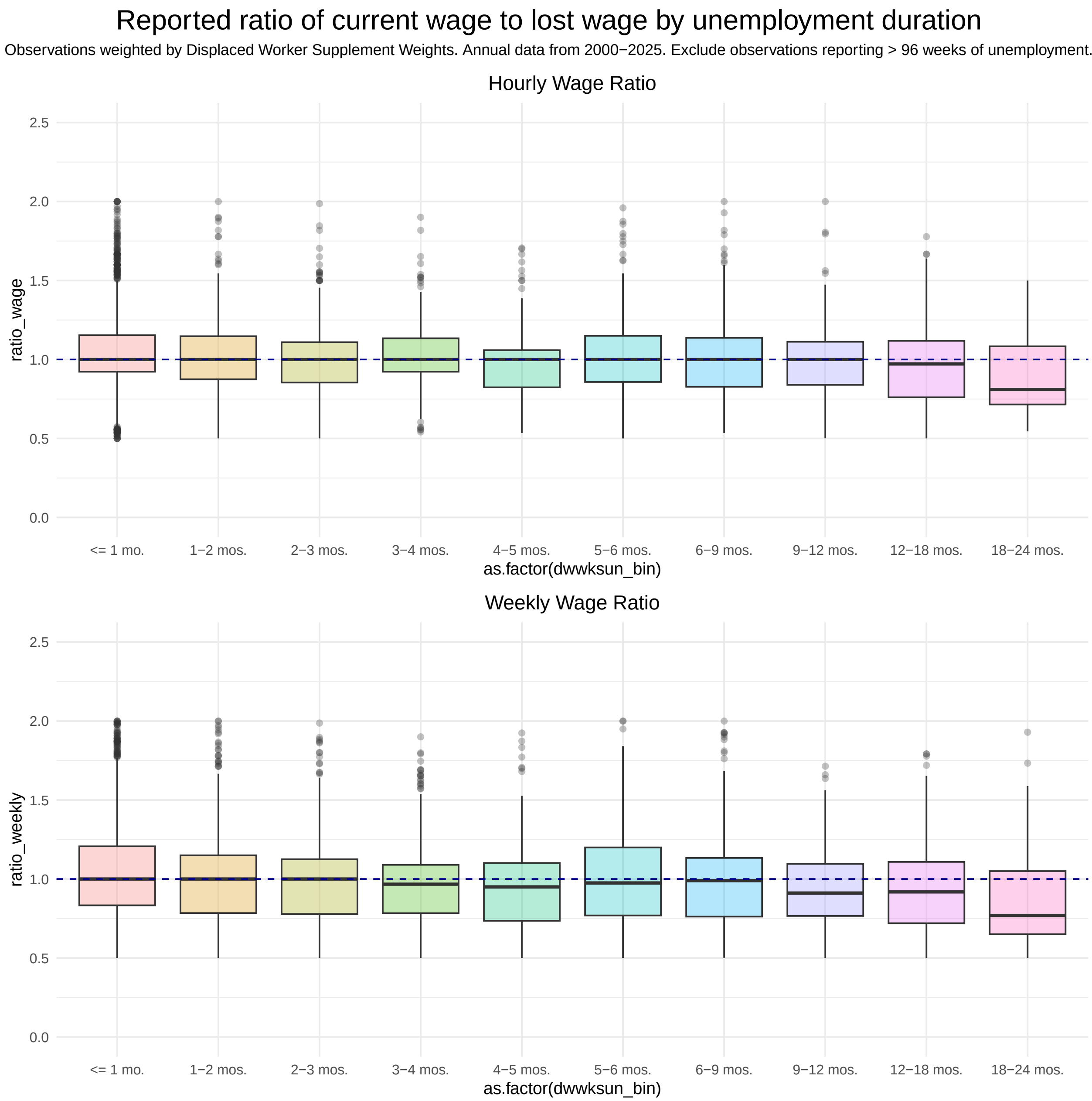
All descriptives below use the Displaced Worker Sample Weights.

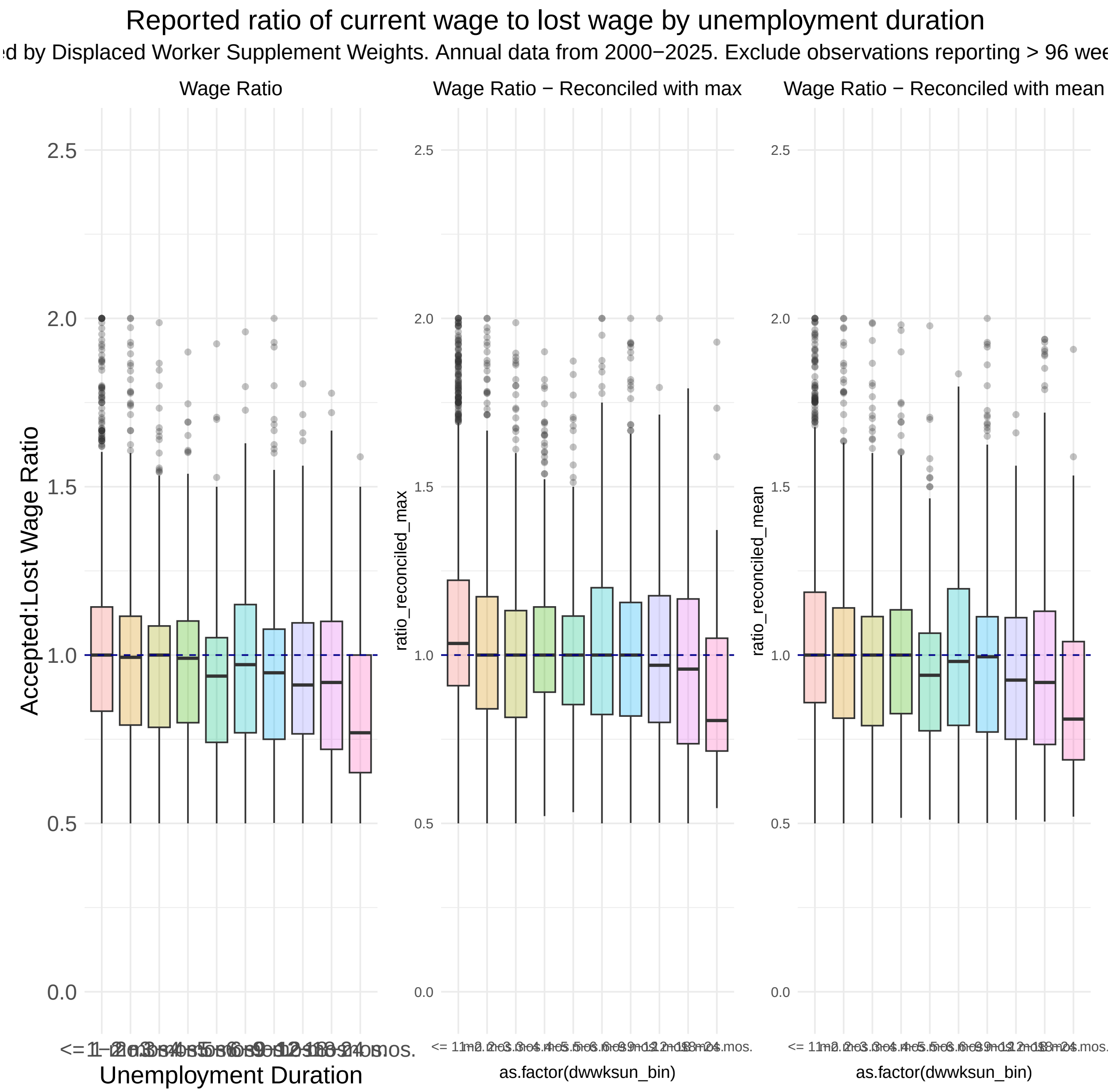
Histogram: sample is skewed (see reweighting alternatives at end of document).

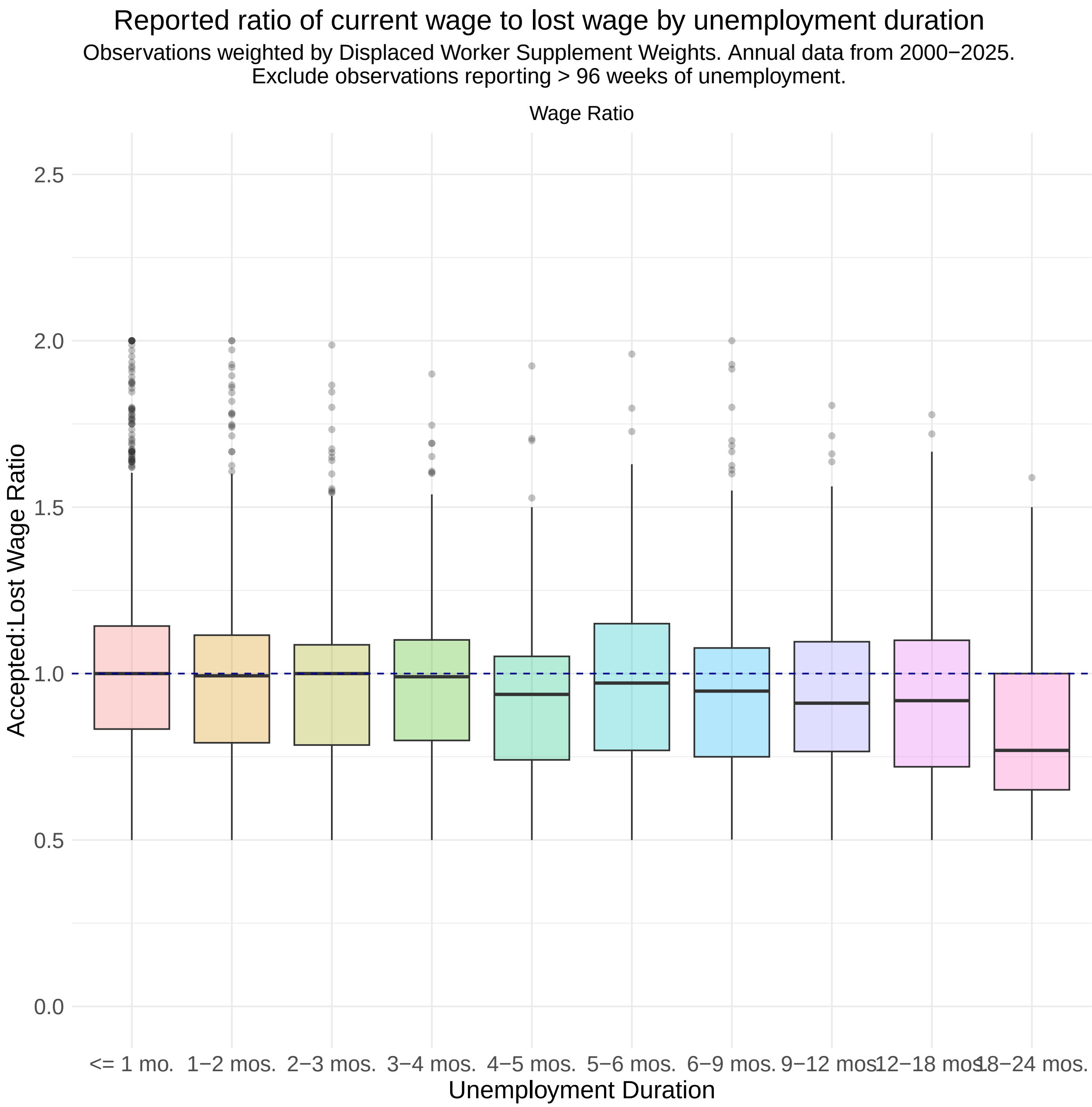
Box plots: Looking at the reported wage ratios in weekly and hourly values, the mean is fixed near 1 until >12 mos of unemployment in hourly wage reporting. In weekly wage reporting, the “satisficing” seems to start earlier in unemployment duration (sample size is larger for weekly reporting - might be worth focusing on those wages).

Scatter plot: I fit a linear and spline fit to the scatted plot of the wage ratio to unemployment duration before using the regression. Indicates decline in the wage ratio with unemployment duration that has a potentially non-linear fit.



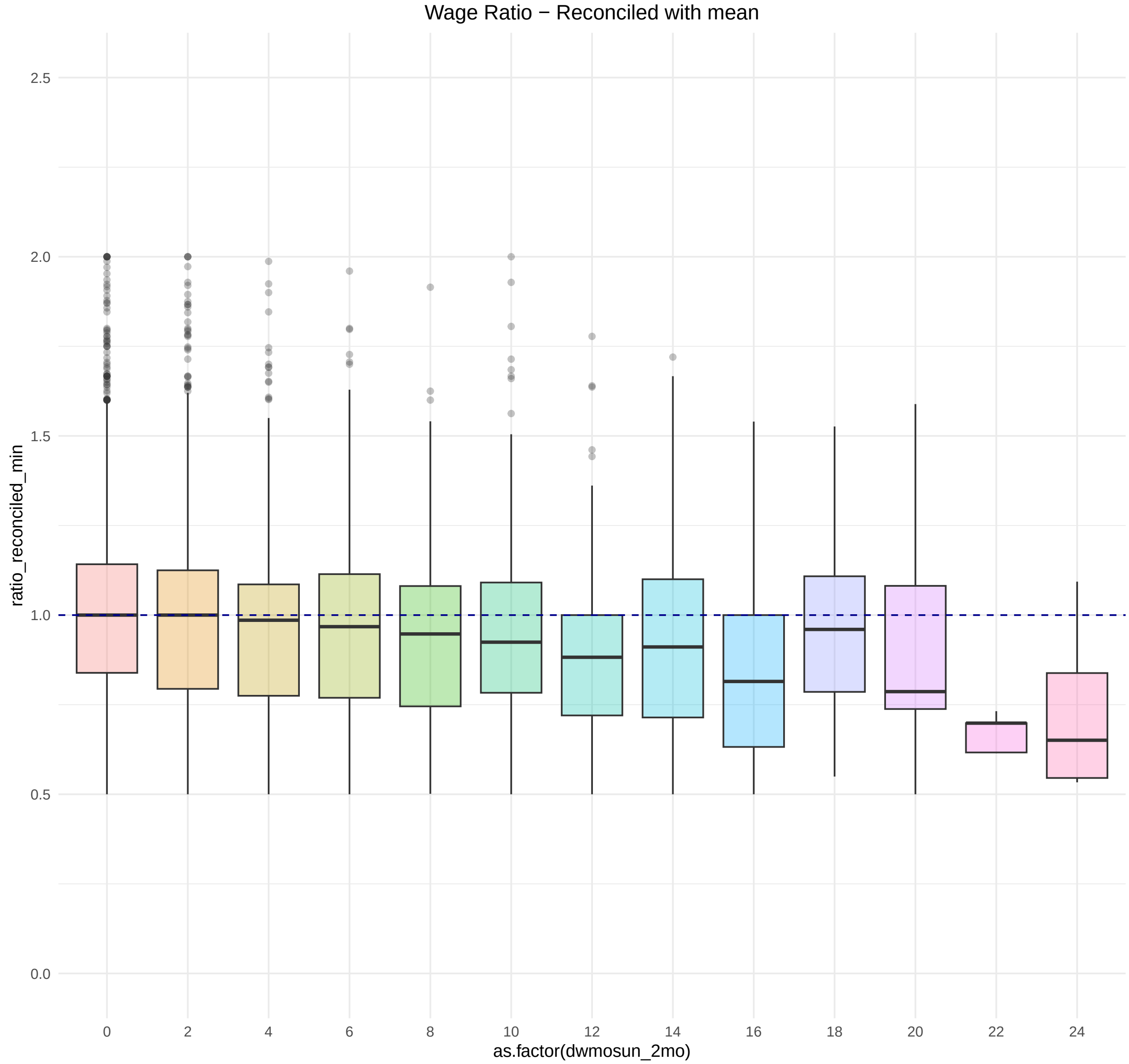

Reported ratio of current wage to lost wage by unemployment duration
Observations weighted by Displaced Worker Supplement Weights. Annual data from 2000−2025. Exclude observations reporting > 96 weeks of unemployment.
Hourly Wage Ratio
ratio_wage
2.5
2.0
1.5
1.0
0.5
0.0
<= 1 mo.
1−2 mos.
2−3 mos.
3−4 mos.
4−5 mos.
5−6 mos.
6−9 mos.
9−12 mos.
12−18 mos.
18−24 mos.
as.factor(dwwksun_bin)
Weekly Wage Ratio
ratio_weekly
2.5
2.0
1.5
1.0
0.5
0.0
<= 1 mo.
1−2 mos.
2−3 mos.
3−4 mos.
4−5 mos.
5−6 mos.
6−9 mos.
9−12 mos.
12−18 mos.
18−24 mos.
as.factor(dwwksun_bin)

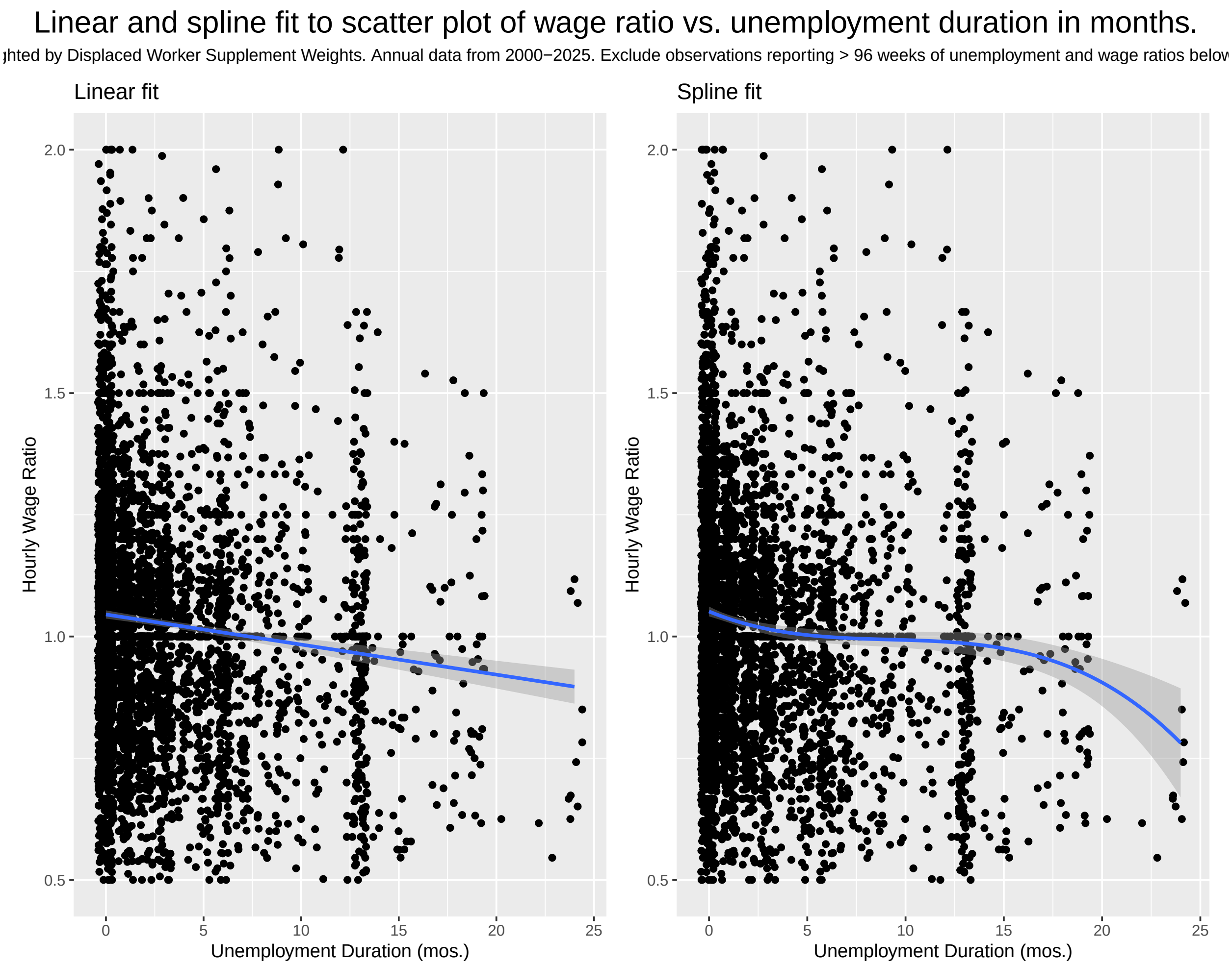

Reported ratio of current wage to lost wage by unemployment duration
:d by Displaced Worker Supplement Weights. Annual data from 2000−2025. Exclude observations reporting > 96 wee
Wage Ratio
Wage Ratio − Reconciled with max
Wage Ratio − Reconciled with mean
Accepted:Lost Wage Ratio
ratio_reconciled_max
ratio_reconciled_mean
2.5
2.0
1.5
1.0
0.5
0.0
Unemployment Duration
as.factor(dwwksun_bin)
as.factor(dwwksun_bin)


Reported ratio of current wage to lost wage by unemployment duration
Observations weighted by Displaced Worker Supplement Weights. Annual data from 2000−2025.
Exclude observations reporting > 96 weeks of unemployment.
Wage Ratio
Accepted:Lost Wage Ratio
2.5
2.0
1.5
1.0
0.5
0.0
<= 1 mo.
1−2 mos.
2−3 mos.
3−4 mos.
4−5 mos.
5−6 mos.
6−9 mos.
9−12 mos.
12−18 mos.
18−24 mos.
Unemployment Duration

Reported ratio of current wage to lost wage by unemployment duration

Observations weighted by Displaced Worker Supplement Weights. Annual data from 2000−2025. Exclude observations reporting > 96 weeks of unemployment.





Regressions (non-uniform sample)

Next, (ignoring for now the non-uniformity of the sample ie. that there are less observations present for higher unemployment durations) I run the following regression (with various modifications to sample and control variables). $W_i = \alpha_i + \beta_1 d_i + \beta_2 UI_i + \beta_3 X_i + \epsilon_i$

where W_i: Ratio of accepted wage to wage at lost job (hourly values).

d_i: Unemployment duration (continuous or binned).

UI_i: Control variable for having used or exhausted unemployment benefits.

X_i: Vector of control variables (sex, age, race (white, black, mixed), marital status (married or not), whether individual used UI benefits, whether individual exhausted UI benefits, education level, and previous wage level).

There are 48 models present with all combinations of the following:

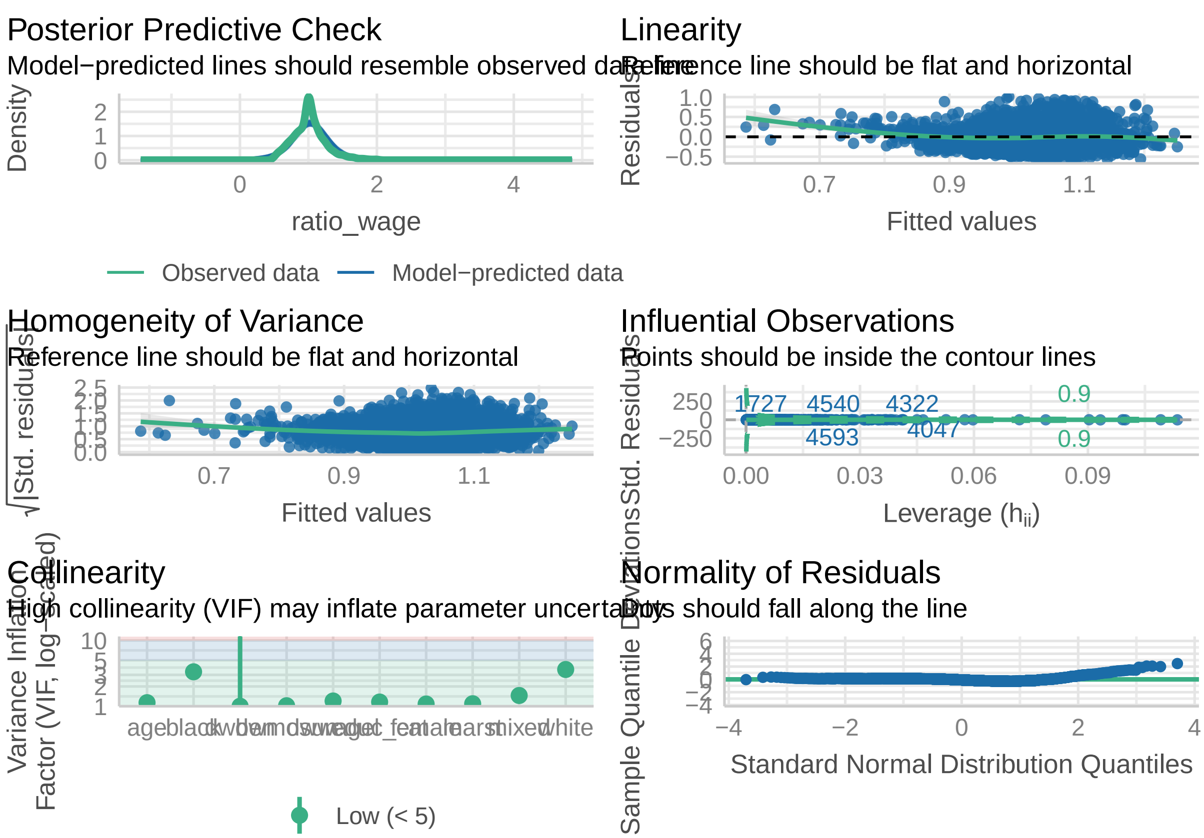
- **Continuous vs. Discrete Treatment Variable (2 alternatives):** Continuous (monthly) versus binned unemployment duration.
- **w. UI vs w. Exhausted UI (3 alternatives):** The data includes a variable for whether individuals USE and/or EXHAUST unemployment benefits. I run the regressions without these UI controls, with control for having used UI, with control for having exhausted UI.

- **w. Controls (2 alternatives):** With or without additional demographic controls (sex, age, race, married, education)
- **w. Wage Level (2 alternatives):** With or without wage level of lost job to control for income. The level of the previous wage likely affects the wage ratio.
- **Outlier clipped sample (2 alternatives):** (As described in the intro section) Remove outliers where the wage ratio is within [0.25, 2.5] and reported unemploymetn duration is below 96 weeks (~ 2 years).

I include the full set of coefficients (again, apologies for verbose output) in case you find the coefficients on the controls interesting (I think the coefficient on age and holding a Bachelor's degree particularly interesting). But I highlight in blue our main interest in β_1.

Across all models in the tabs below we see a consistently negative coefficient on unemployment duration (~0.7-1 percentage point increase in the wage ratio for each additional month spent in unemployment). If we look more closely at the performance of our model with continuous unemployment duration, UI use (not exhaustion), all controls, wage levels, and outlier correction we see that the model performs fairly well across various diagnostic tests.

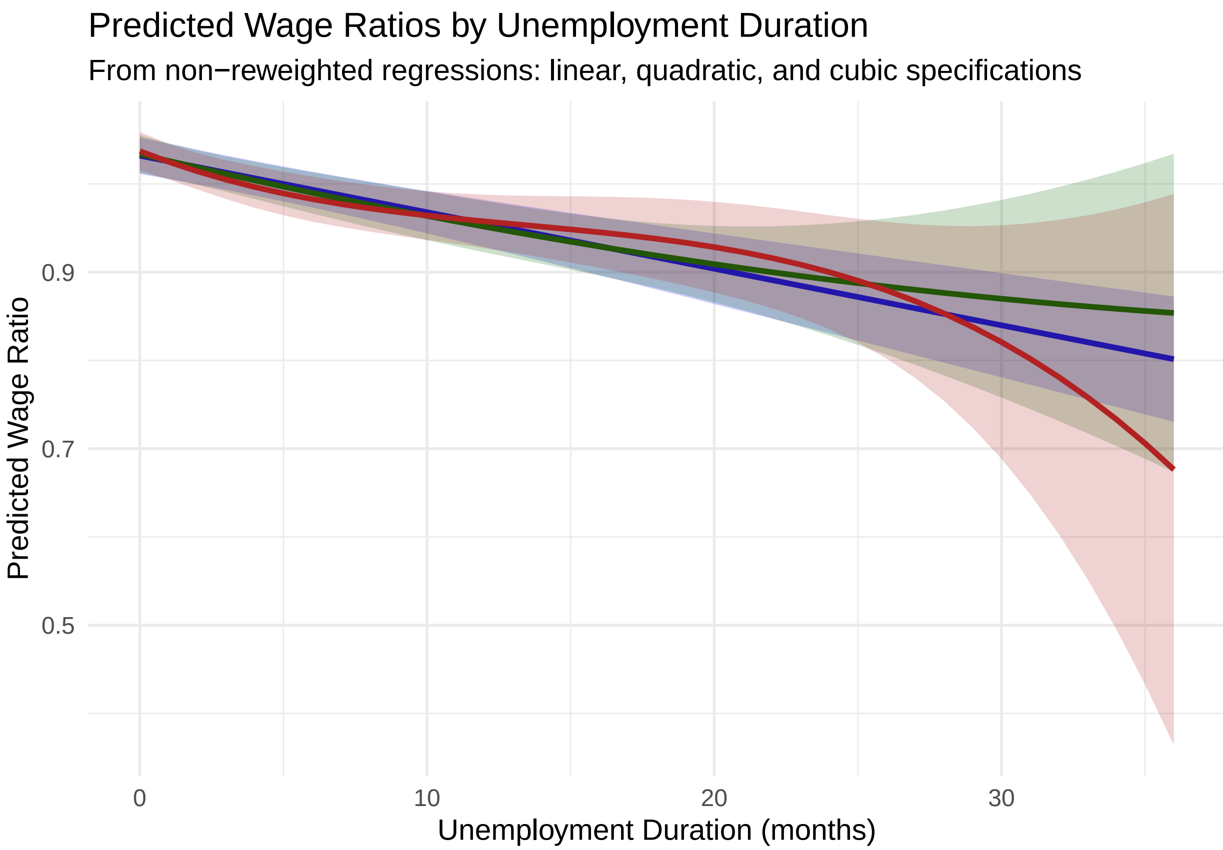
```
## [1] "Continuous U Duration. w. UI Control w. demographic controls (clipped sample)"
```



Continuous UE Duration

Continuous UE duration treatment is reported in monthly values. A one-unit increase in the treatment variable = 1 additional month of unemployment.

W.O. Wage Level Control



W. Wage Level Control

Binned UE Duration

Binned UE duration treatment is reported in bins as indicated in the box plots and code cleaning above.

W.O. Wage Level Control

W. Wage Level Control

Additional Considerations

Below I:

1. Show results from some sample reweighting to address the non-uniformity of our cross-sectional data.
2. Histogram of an additional explanatory variable that might be interesting - the tenure of the lost job. How long (in years) did the individual hold the job they lost.
3. Additionally, I show some rough graphs/figures about the sample population (age, education, gender, wage distributions). I am still working on an occupational distribution graph to understand the "skills"/occupational distribution of the sample.

Selection Issues with Non-Random Sample

NOTE: Skip ahead to “Regression Results with Sample Reweighting for Regression Results if you don’t wish to look at the reweighting details below.

One of the challenges with this data is that the sample grows significantly smaller for higher reported of unemployment duration (see scatter plots in Descriptives section). One option is a sample reweighting (beyond the census weights) to ensure population similarity across bins (below I choose GLM propensity score matching & entropy-balancing) or a Heckman Selection. Again, I include the code below (apologies for verbose output), mainly because I am not yet 100% sure of the implementation as I have never implemented such sample correction in a cross-sectional study). Open to suggestions and corrections :)

Conclusion: With this implementation (which may very well be wrong for now!), the coefficients on unemployment duration remain stable.

Entropy Balancing Entropy balancing simply reweights observations to ensure population matching across the key dependent variable.

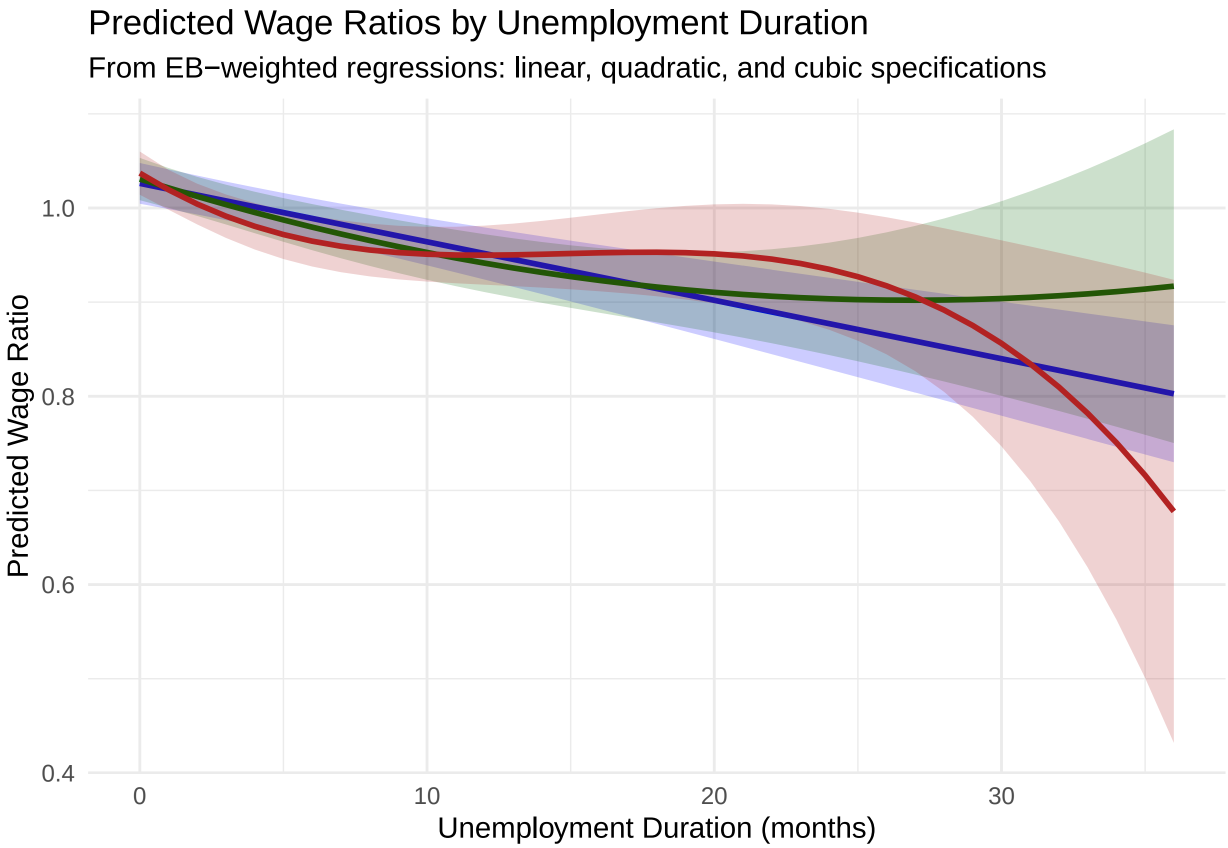
```r
# Apply entropy balancing using dwsuppwt sample weights
# Reweight according to observable characteristics using "ebalance"
eb <- weightit(
  formula = dwmosun ~ female + age + white + black + mixed + marst + educ_cat,
  data = df,
  method = "ebalance",
  s.weights = df$dwsuppwt
)

# All covariates are balanced at the mean with tight threshold
bal.tab(eb, stats = c("m", "v"), thresholds = c(m = .001))
```

```
## Balance Measures
##                                 Type Diff.Target.Adj      M.Threshold
## female                        Binary          0.0000 Balanced, <0.001
## age                          Contin.         -0.0000 Balanced, <0.001
## white                         Binary         -0.0000 Balanced, <0.001
## black                         Binary         -0.0000 Balanced, <0.001
## mixed                         Binary          0.0001 Balanced, <0.001
## marst                         Binary         -0.0000 Balanced, <0.001
## educ_cat_Less than HS         Binary         -0.0000 Balanced, <0.001
## educ_cat_HS Diploma           Binary         -0.0000 Balanced, <0.001
## educ_cat_Associate's          Binary         -0.0000 Balanced, <0.001
## educ_cat_Bachelor's           Binary         -0.0000 Balanced, <0.001
## educ_cat_Postgraduate Degree  Binary          0.0001 Balanced, <0.001
##
## Balance tally for target mean differences
##                       count
## Balanced, <0.001         11
## Not Balanced, >0.001      0
##
## Variable with the greatest target mean difference
##  Variable Diff.Target.Adj      M.Threshold
##     mixed          0.0001 Balanced, <0.001
##
## Effective sample sizes
```

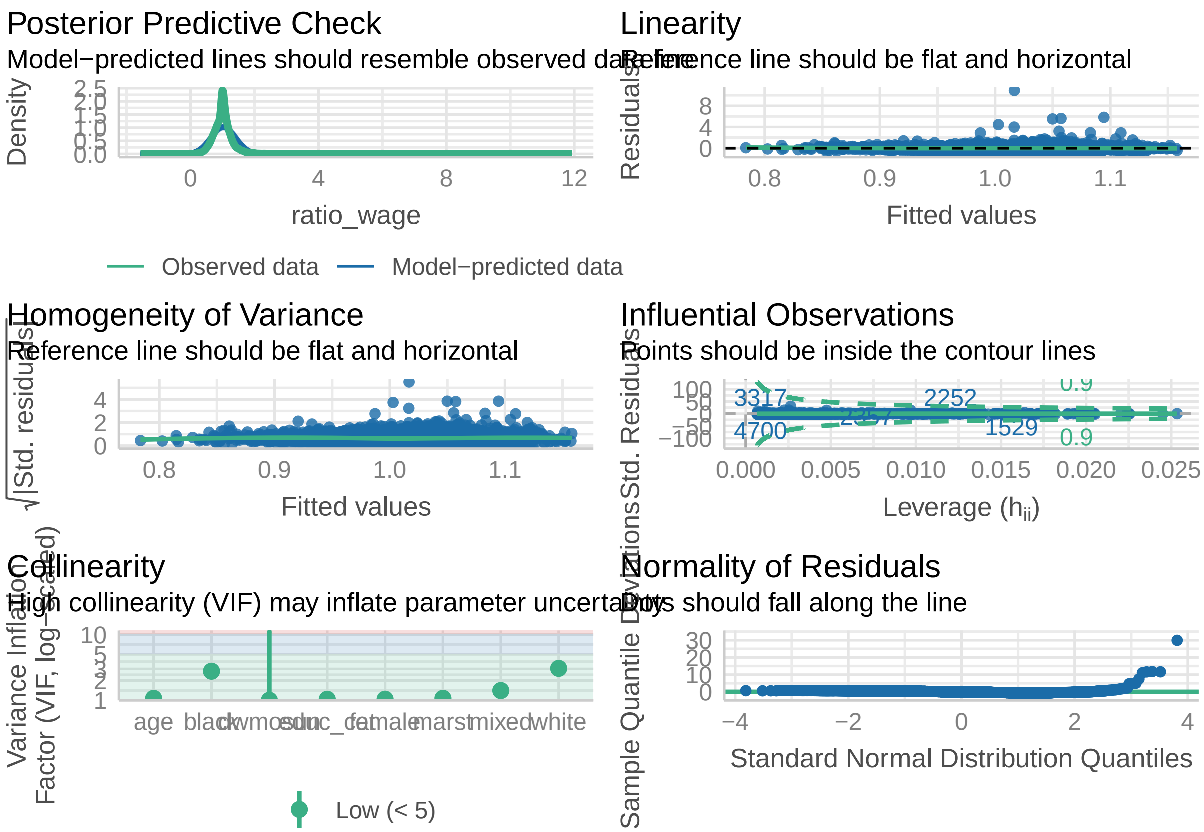
```
##              Total
## Unadjusted 4747.86
## Adjusted   4634.14
```

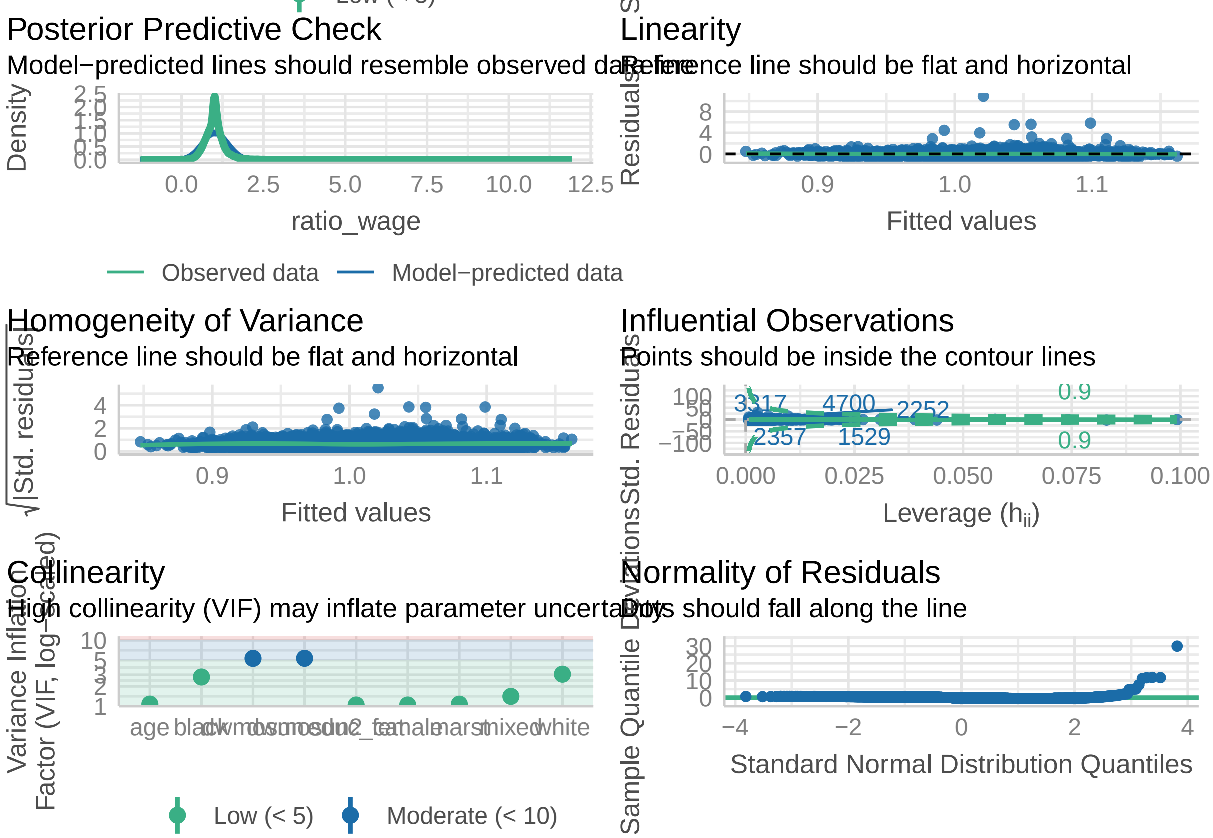
```
# Add the new weights to the dataframe
df$eb_weight <- eb$weights

# Run weighted linear regression using entropy-balanced weights
mod_eb_reweight <- lm(
  formula = ratio_wage ~ dwmosun + female + age + white + black + mixed + marst + educ_cat,
  data = df,
  weights = eb_weight
)

mod_eb_reweight2 <- lm(
  formula = ratio_wage ~ dwmosun + dwmosun2 + female + age + white + black + mixed + marst + educ_cat,
  data = df,
  weights = eb_weight
)

mod_eb_reweight3 <- lm(
  formula = ratio_wage ~ dwmosun + dwmosun2 + dwmosun3 + female + age + white + black + mixed + marst +
  data = df,
  weights = eb_weight
)


# Extract prediction intervals as data frames
eb_preds  <- as.data.frame(predict(mod_eb_reweight,  newdata = newdata, interval = "confidence")) %>% re
eb_preds2 <- as.data.frame(predict(mod_eb_reweight2, newdata = newdata, interval = "confidence")) %>% re
eb_preds3 <- as.data.frame(predict(mod_eb_reweight3, newdata = newdata, interval = "confidence")) %>% re

eb_full <- cbind(eb_preds, eb_preds2, eb_preds3)

eb_full %>%
  ggplot(aes(x = 0:(nrow(eb_full)-1))) +
    # Model 1
    geom_line(aes(y = eb_preds_fit), color = "blue", linewidth = 1) +
    geom_ribbon(aes(ymin = eb_preds_lwr, ymax = eb_preds_upr), fill = "blue", alpha = 0.2) +

    # Model 2
    geom_line(aes(y = eb_preds2_fit), color = "darkgreen", linewidth = 1) +
    geom_ribbon(aes(ymin = eb_preds2_lwr, ymax = eb_preds2_upr), fill = "darkgreen", alpha = 0.2) +

    # Model 3
    geom_line(aes(y = eb_preds3_fit), color = "firebrick", linewidth = 1) +
    geom_ribbon(aes(ymin = eb_preds3_lwr, ymax = eb_preds3_upr), fill = "firebrick", alpha = 0.2) +

    labs(
      title = "Predicted Wage Ratios by Unemployment Duration",
      subtitle = "From EB-weighted regressions: linear, quadratic, and cubic specifications",
      x = "Unemployment Duration (months)",
      y = "Predicted Wage Ratio"
    ) +
```

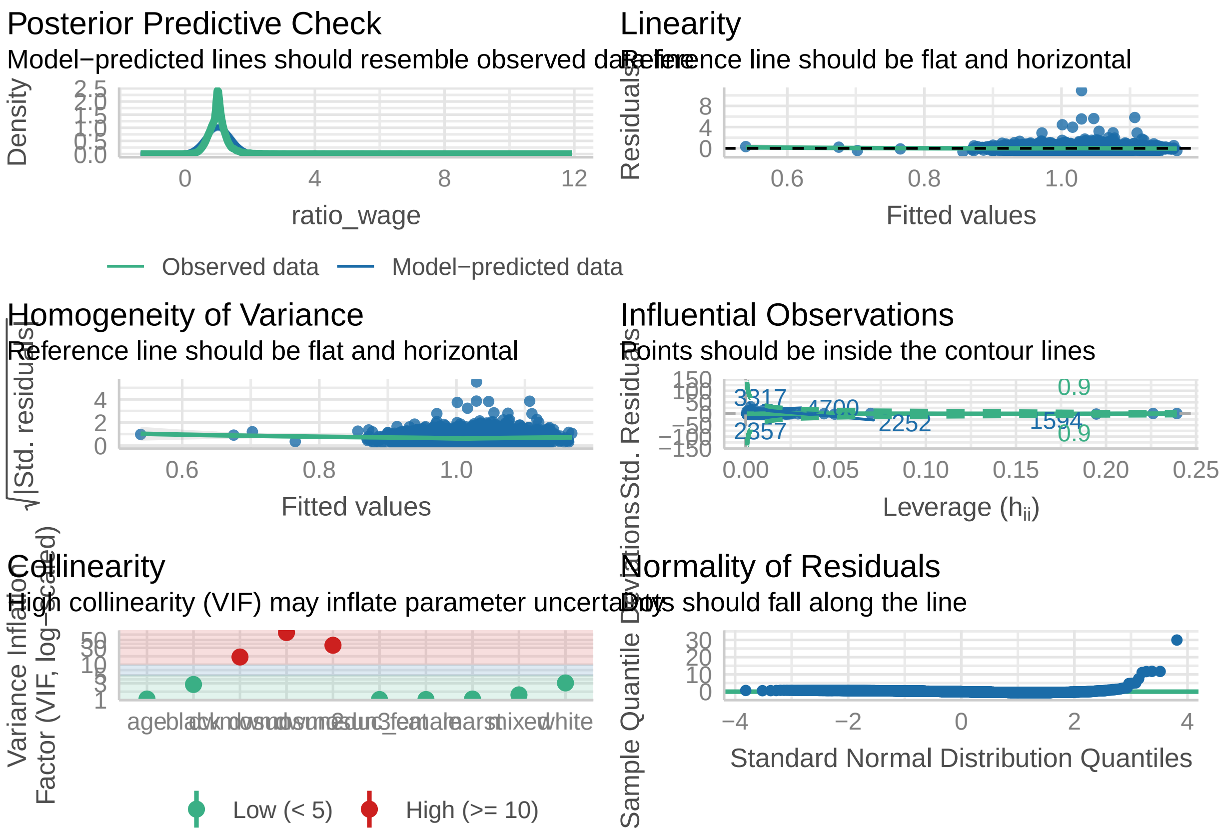
```r
    theme_minimal()
```




```
## [1] "Diagnostic Tests for Entropy-balanced Reweighted Sample"
```


Posterior Predictive Check
Model−predicted lines should resemble observed data
Density
ratio_wage
Observed data
Model−predicted data
Linearity
Reference line should be flat and horizontal
Residuals
Fitted values
Homogeneity of Variance
Reference line should be flat and horizontal
√|Std. residuals|
Fitted values
Influential Observations
Points should be inside the contour lines
Std. Residuals
Leverage (h_{ii})
Collinearity
High collinearity (VIF) may inflate parameter uncertainty
Variance Inflation Factor (VIF, log−scaled)
age
black
dwmosdur
educ_cat
female
marst
mixed
white
Low (< 5)
Normality of Residuals
Dots should fall along the line
Sample Quantile Deviations
Standard Normal Distribution Quantiles


Posterior Predictive Check
Model−predicted lines should resemble observed data
Density
ratio_wage
Observed data
Model−predicted data
Linearity
Reference line should be flat and horizontal
Residuals
Fitted values
Homogeneity of Variance
Reference line should be flat and horizontal
√|Std. residuals|
Fitted values
Influential Observations
Points should be inside the contour lines
Std. Residuals
Leverage (h_{ii})
Collinearity
High collinearity (VIF) may inflate parameter uncertainty
Variance Inflation Factor (VIF, log−scaled)
Low (< 5)
Moderate (< 10)
Normality of Residuals
Dots should fall along the line
Sample Quantile Deviations
Standard Normal Distribution Quantiles




```
# More conventional propensity scoring with a GLM estimator
glm <- weightit(
  formula = dwmosun ~ female + age + white + black + mixed + marst + educ_cat,
  data = df,
  method = "glm",
  s.weights = df$dwsuppwt
)

# All covariates are balanced at the mean with less tight threshold (0.001, very few variables pass wit
bal.tab(glm, stats = c("m", "v"), thresholds = c(m = .01))
```

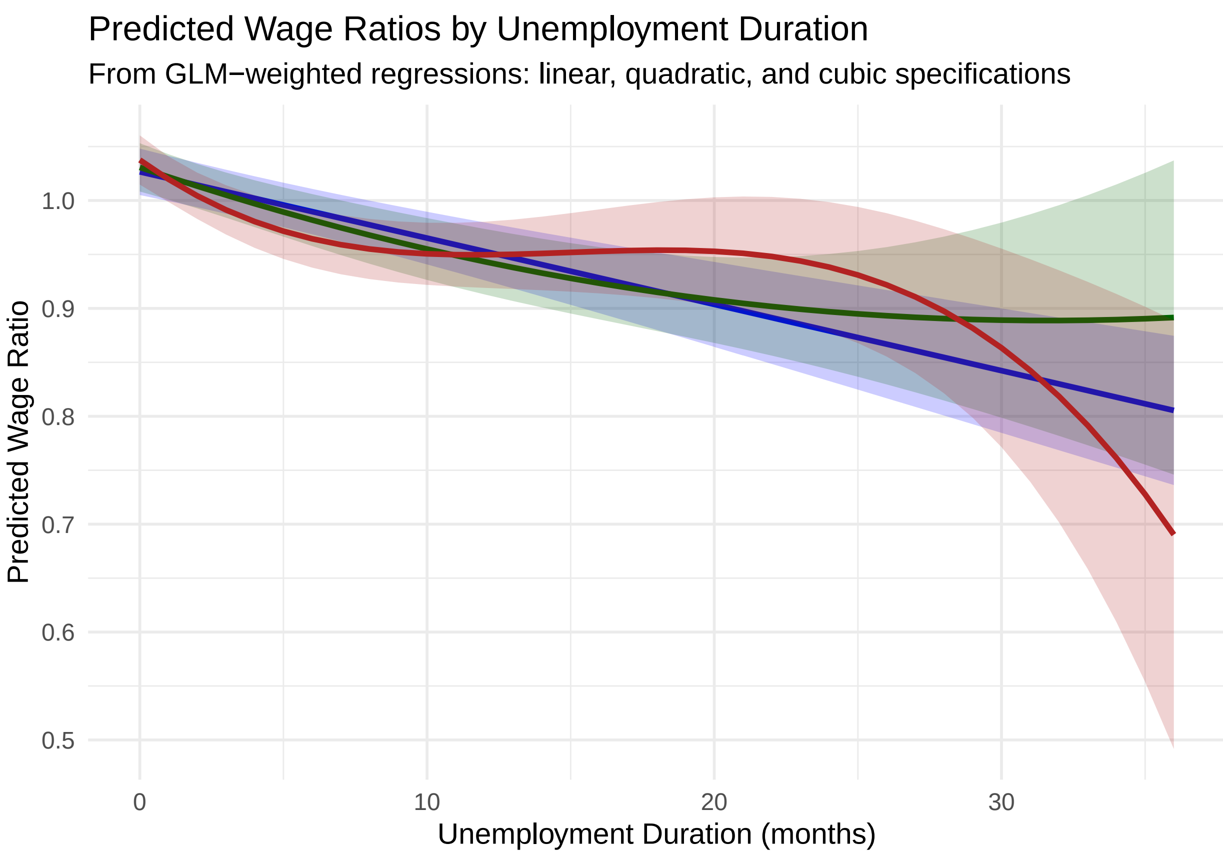
Propensity Score Weighting with GLM Estimator

```
## Balance Measures
##                                  Type Diff.Target.Adj     M.Threshold
## female                         Binary         -0.0032 Balanced, <0.01
## age                           Contin.         -0.0057 Balanced, <0.01
## white                          Binary          0.0022 Balanced, <0.01
## black                          Binary         -0.0005 Balanced, <0.01
## mixed                          Binary         -0.0041 Balanced, <0.01
## marst                          Binary          0.0041 Balanced, <0.01
## educ_cat_Less than HS          Binary         -0.0042 Balanced, <0.01
## educ_cat_HS Diploma            Binary         -0.0003 Balanced, <0.01
## educ_cat_Associate's           Binary          0.0032 Balanced, <0.01
```

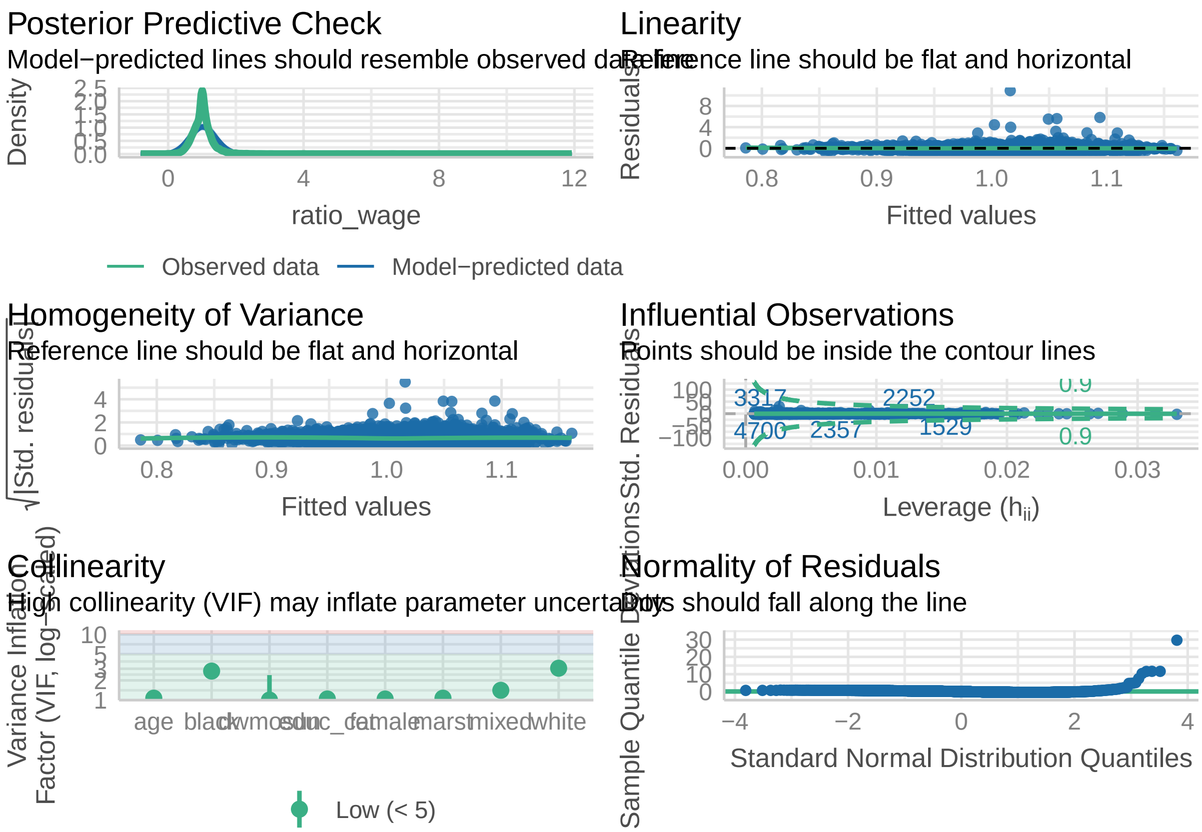
```
## educ_cat_Bachelor's           Binary          0.0023 Balanced, <0.01
## educ_cat_Postgraduate Degree  Binary         -0.0017 Balanced, <0.01
##
## Balance tally for target mean differences
##                     count
## Balanced, <0.01        11
## Not Balanced, >0.01     0
##
## Variable with the greatest target mean difference
##  Variable Diff.Target.Adj     M.Threshold
##       age         -0.0057 Balanced, <0.01
##
## Effective sample sizes
##              Total
## Unadjusted 4747.86
## Adjusted   4637.07
```

```
df$glm_weight <- glm$weights

mod_glm_reweight <- lm(
  formula = ratio_wage ~ dwmosun + female + age + white + black + mixed + marst + educ_cat,
  data = df,
  weights = glm_weight
)
mod_glm_reweight2 <- lm(
  formula = ratio_wage ~ dwmosun + dwmosun2 + female + age + white + black + mixed + marst + educ_cat,
  data = df,
  weights = glm_weight
)
mod_glm_reweight3 <- lm(
  formula = ratio_wage ~ dwmosun + dwmosun2 + dwmosun3 + female + age + white + black + mixed + marst +
  data = df,
  weights = glm_weight
)

# Extract prediction intervals as data frames
glm_preds  <- as.data.frame(predict(mod_glm_reweight,  newdata = newdata, interval = "confidence")) %>%
glm_preds2 <- as.data.frame(predict(mod_glm_reweight2, newdata = newdata, interval = "confidence")) %>%
glm_preds3 <- as.data.frame(predict(mod_glm_reweight3, newdata = newdata, interval = "confidence")) %>%

glm_full <- cbind(glm_preds, glm_preds2, glm_preds3)

glm_full %>%
  ggplot(aes(x = 0:36)) +
    # Model 1
    geom_line(aes(y = glm_preds_fit), color = "blue", linewidth = 1) +
    geom_ribbon(aes(ymin = glm_preds_lwr, ymax = glm_preds_upr), fill = "blue", alpha = 0.2) +

    # Model 2
    geom_line(aes(y = glm_preds2_fit), color = "darkgreen", linewidth = 1) +
    geom_ribbon(aes(ymin = glm_preds2_lwr, ymax = glm_preds2_upr), fill = "darkgreen", alpha = 0.2) +

    # Model 3
    geom_line(aes(y = glm_preds3_fit), color = "firebrick", linewidth = 1) +
```

```
    geom_ribbon(aes(ymin = glm_preds3_lwr, ymax = glm_preds3_upr), fill = "firebrick", alpha = 0.2) +

    labs(
      title = "Predicted Wage Ratios by Unemployment Duration",
      subtitle = "From GLM-weighted regressions: linear, quadratic, and cubic specifications",
      x = "Unemployment Duration (months)",
      y = "Predicted Wage Ratio"
    ) +
    theme_minimal()
```



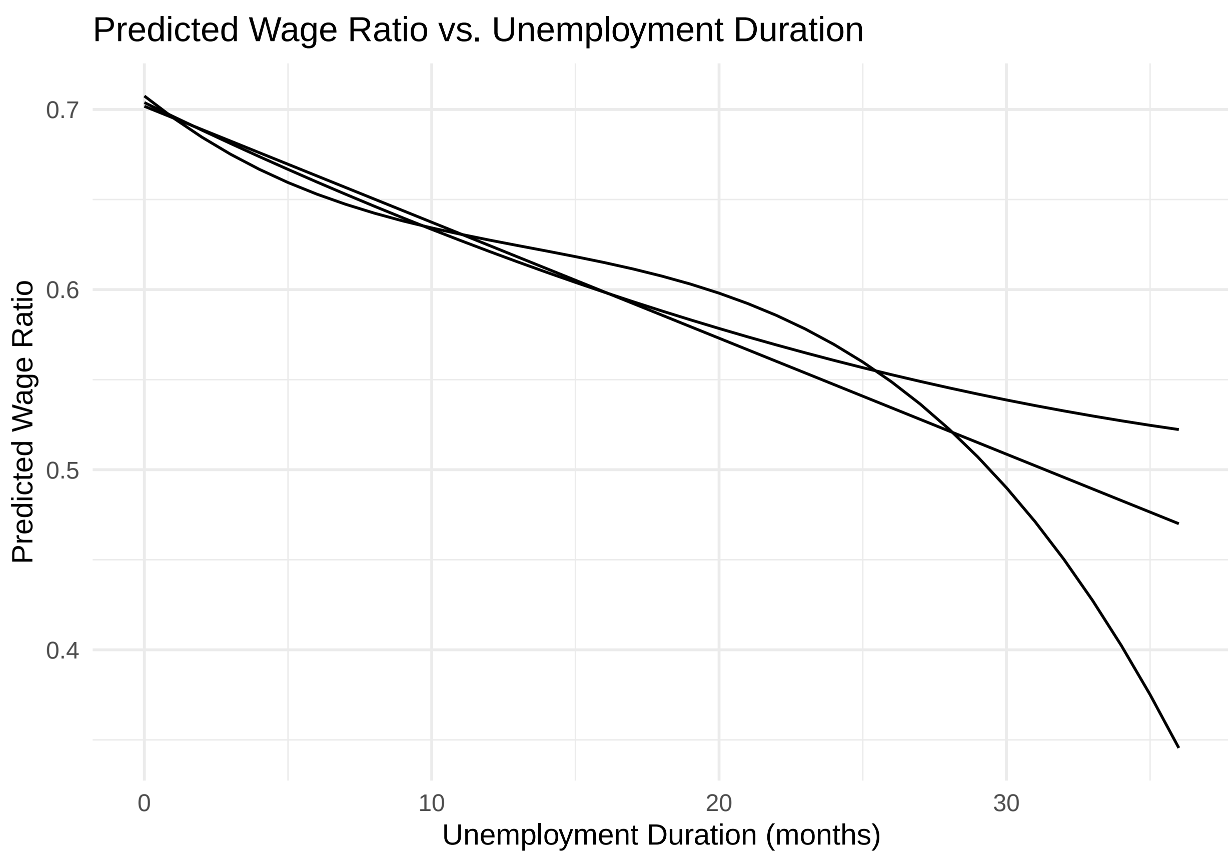

```
## [1] "Diagnostic Tests for Propensity Score Matching (GLM) Reweighted Sample"
```

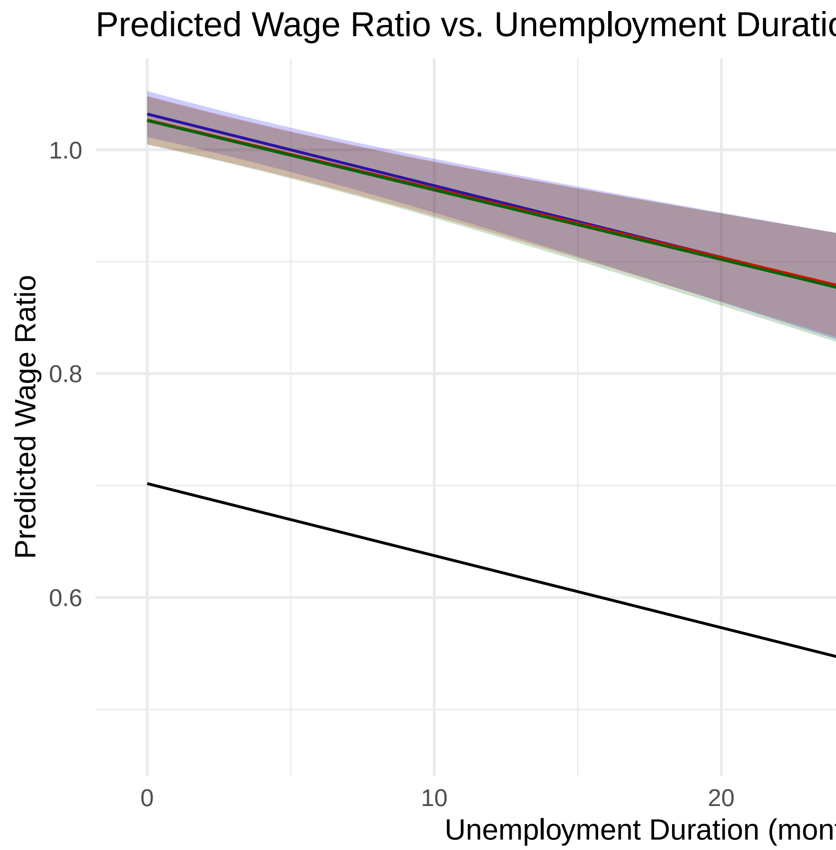


Heckman Selection Another option is a Heckman Selection correction though I do not think this addresses the particular selection concern we have where there are simply less observations in longer unemployment durations.

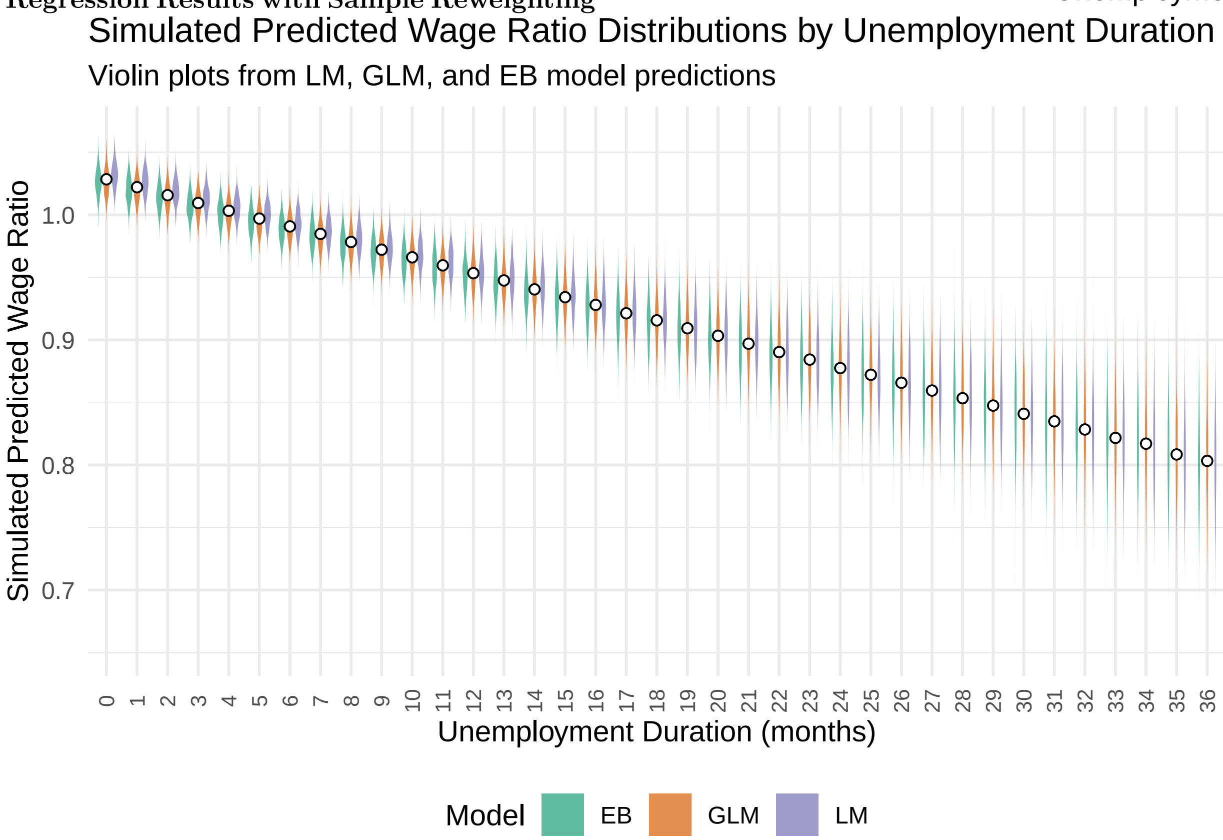
```
# Create selection indicator (1 if ratio_wage is observed)
df$observe_wage <- as.numeric(!is.na(df$ratio_wage))

# Define selection and outcome equations
selection_eq <- observe_wage ~ female + age + white + black + mixed + marst + educ_cat
outcome_eq  <- ratio_wage ~ dwmosun + female + age + white + black + mixed + marst + educ_cat
outcome_eq2 <- ratio_wage ~ dwmosun + dwmosun2 + female + age + white + black + mixed + marst + educ_ca
outcome_eq3 <- ratio_wage ~ dwmosun + dwmosun2 + dwmosun3 + female + age + white + black + mixed + mars

# Run Heckman
heckman_model <- selection(
  selection = selection_eq,
  outcome = outcome_eq,
  data = df,
  method = "2step",
  weights = df$dwsuppwt    # Include weights from CPS
)

# Run Heckman
heckman_model2 <- selection(
  selection = selection_eq,
  outcome = outcome_eq2,
```

```r
  data = df,
  method = "2step",
  weights = df$dwsuppwt   # Include weights from CPS
)

# Run Heckman
heckman_model3 <- selection(
  selection = selection_eq,
  outcome = outcome_eq3,
  data = df,
  method = "2step",
  weights = df$dwsuppwt   # Include weights from CPS
)

# Predict ratio_wage (corrected for selection)
preds <- predict(heckman_model, newdata = newdata, part = "outcome", type = "unconditional")
preds2 <- predict(heckman_model2, newdata = newdata, part = "outcome", type = "unconditional")
preds3 <- predict(heckman_model3, newdata = newdata, part = "outcome", type = "unconditional")


# Combine with original durations for plotting or inspection
newdata$predicted_ratio <- preds
newdata$predicted_ratio2 <- preds2
newdata$predicted_ratio3 <- preds3

ggplot(newdata, aes(x = dwmosun)) +
  geom_line(aes(y = predicted_ratio)) +
  geom_line(aes(y = predicted_ratio2)) +
  geom_line(aes(y = predicted_ratio3)) +
  labs(x = "Unemployment Duration (months)", y = "Predicted Wage Ratio", title = "Predicted Wage Ratio v
  theme_minimal()
```


Predicted Wage Ratio vs. Unemployment Duration
Predicted Wage Ratio
0.7
0.6
0.5
0.4
0
10
20
30
Unemployment Duration (months)

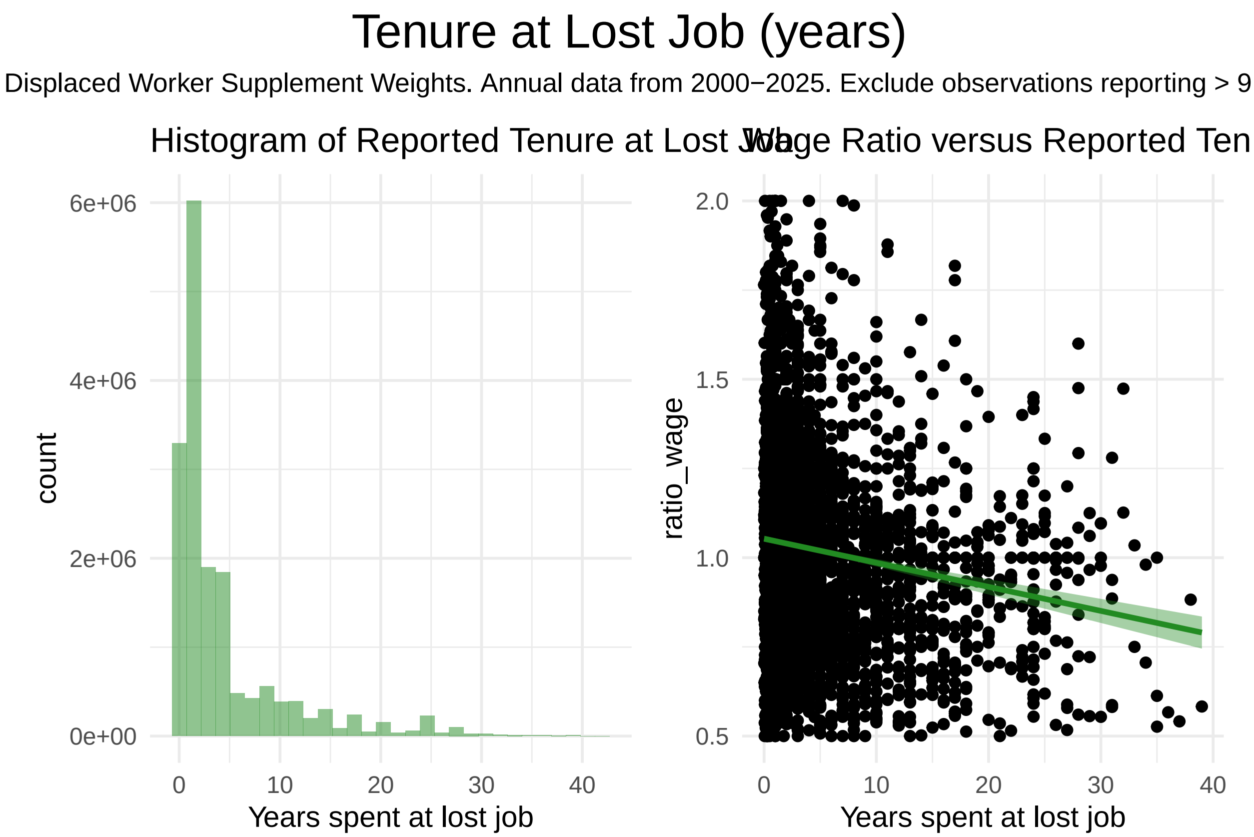


Regression Results with Sample Reweighting



Job Tenure

We have information on the tenure spent at the last job which could impact the result. This could speak to the "adaptability" of individuals. Wage ratio seems to decrease (although not sure if meaningfully) with tenure at previous job.

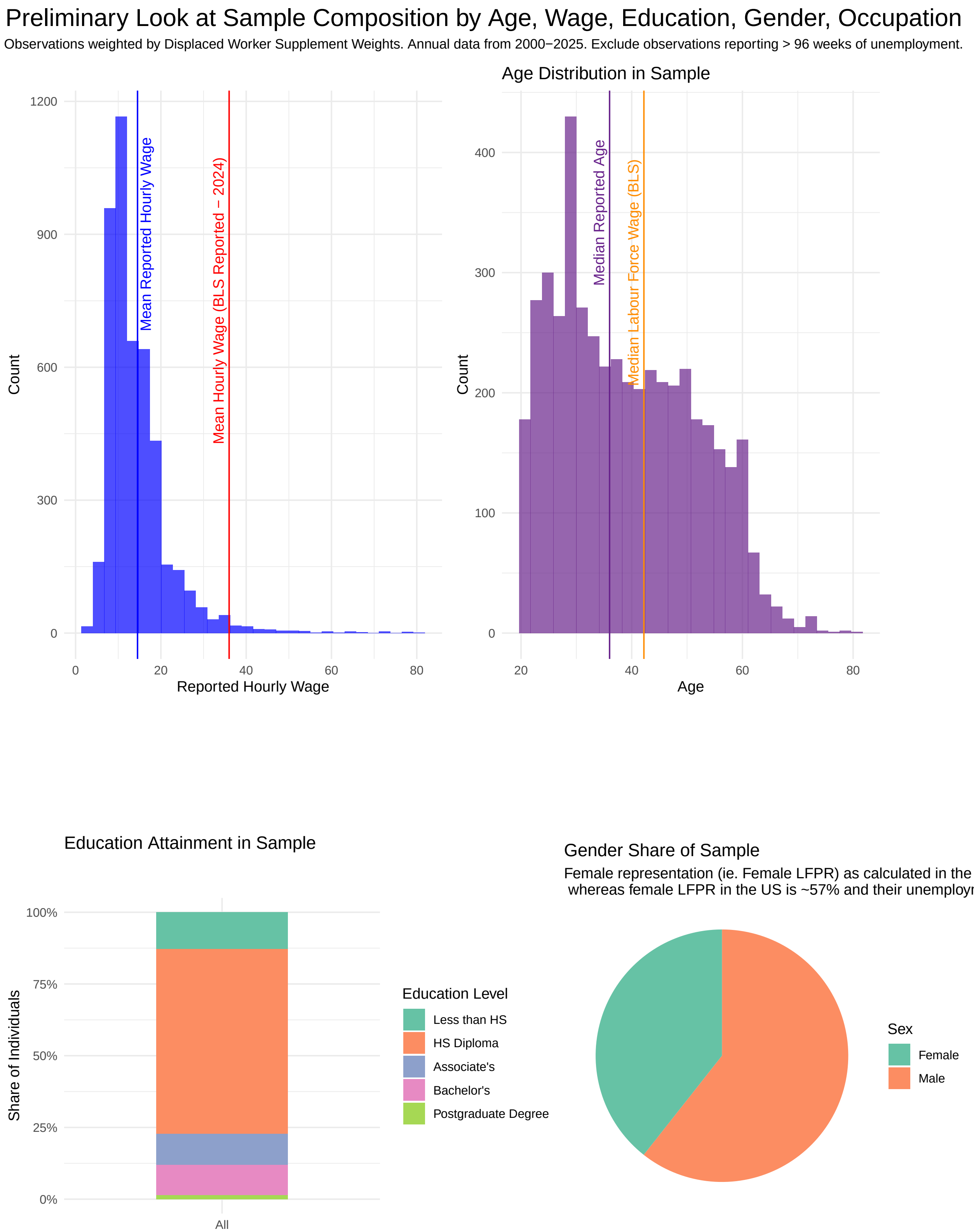


Representation

Although the survey does provide sample weights which we use above, it's still likely that those who are laid off might be systematically more susceptible to layoffs (lower-wage, low-skill occupation, male, etc). Below, some (very rough) graphs to indicate what the sample looks like.

Headline result: it seems the sample over-represents below-mean wage earners and women. Age looks reasonably accurate (in relation to a simple median though....have not checked spread). Have not yet checked match to educational attainment. Individuals with only a HS diploma is strong majority in sample - not sure how accurate this is (likely correlated with wage however...so this might be cause for concern and confirm a skewed sample in that sense).

If we wish to pursue this data, I could improve on the below but it will have to do for now.



OTJ Search

Eeckhout et al. 2019 Unemployment Cycles Source

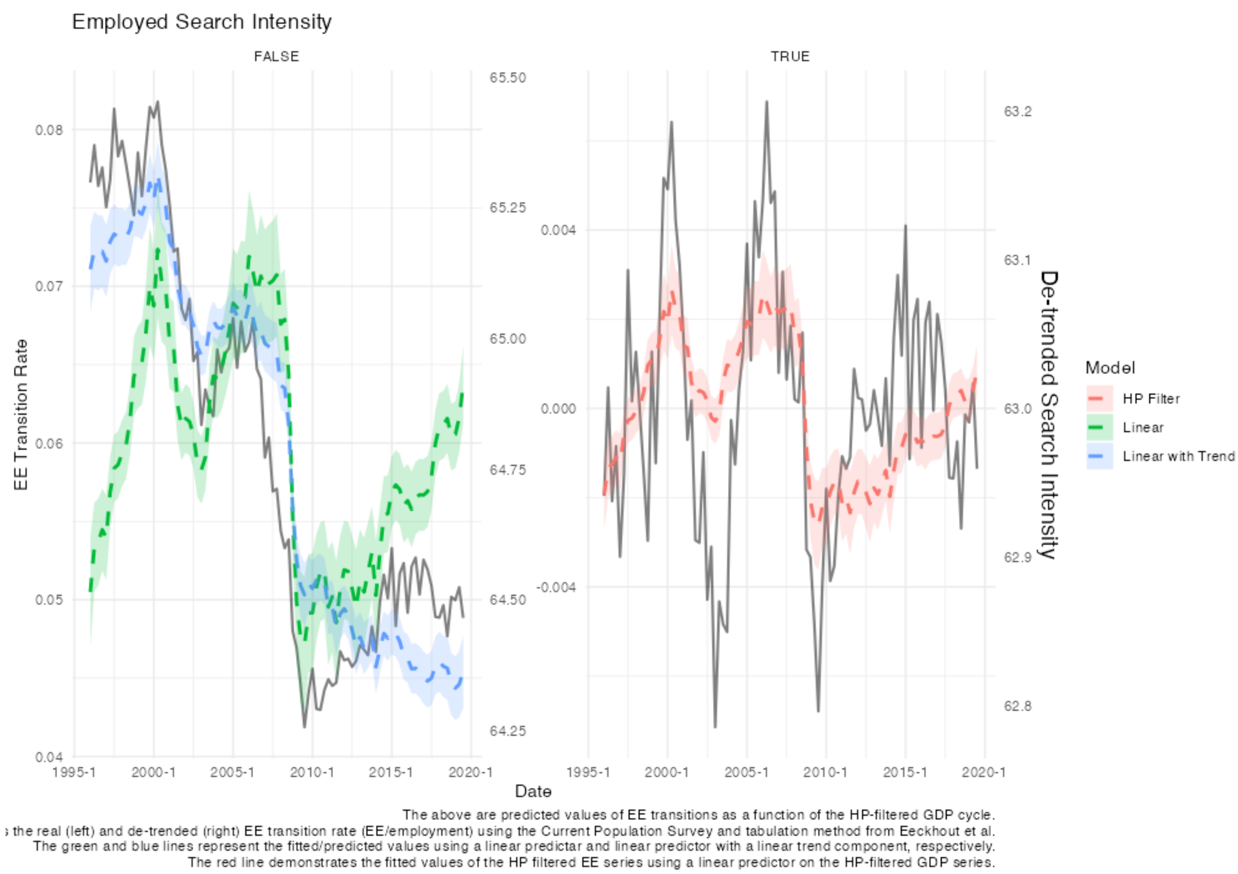
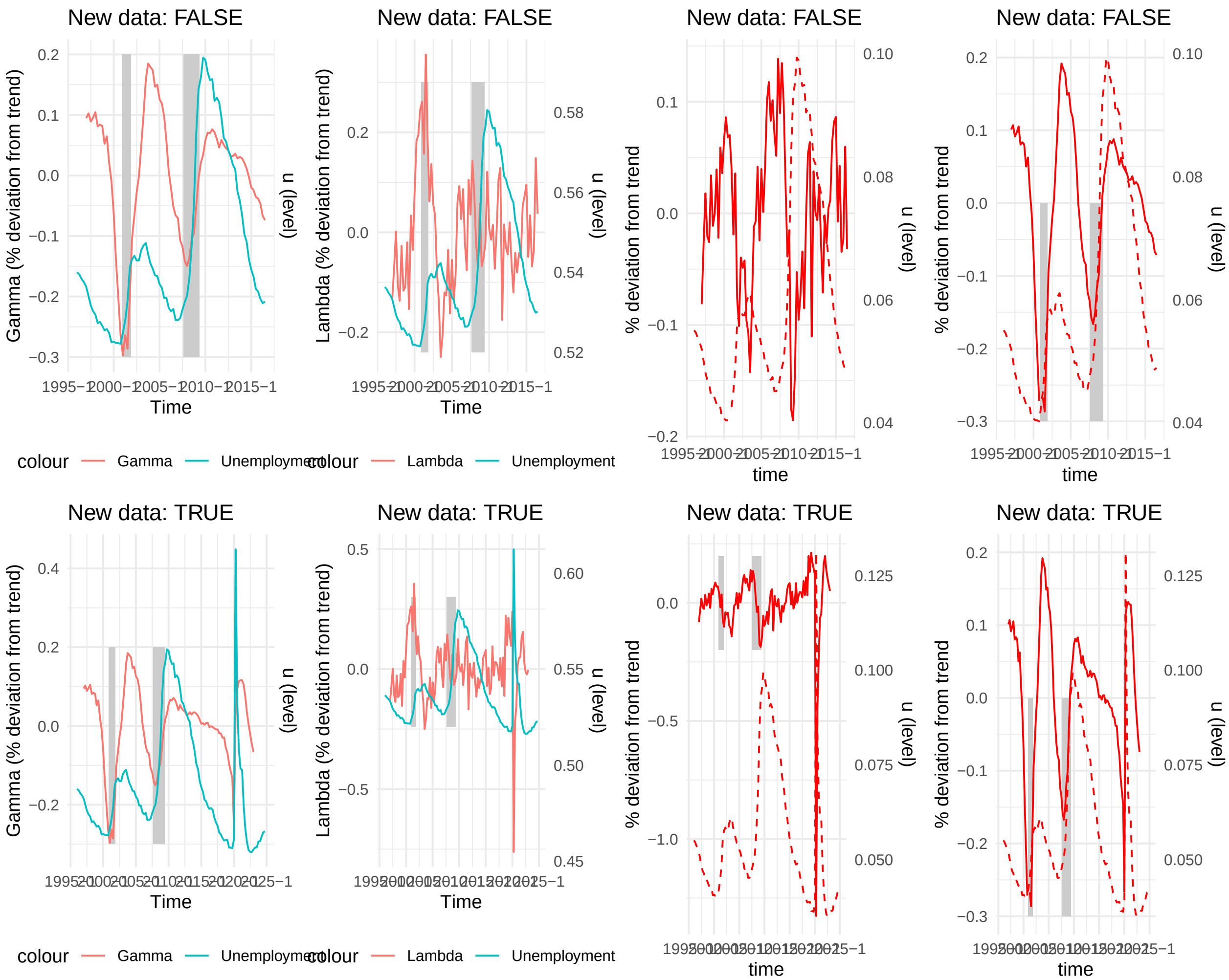
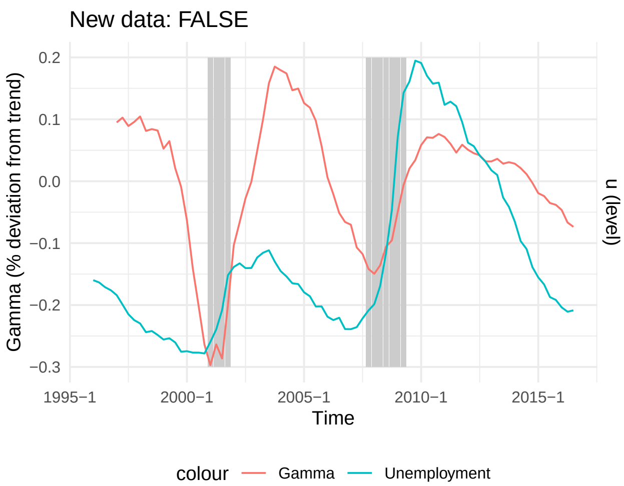
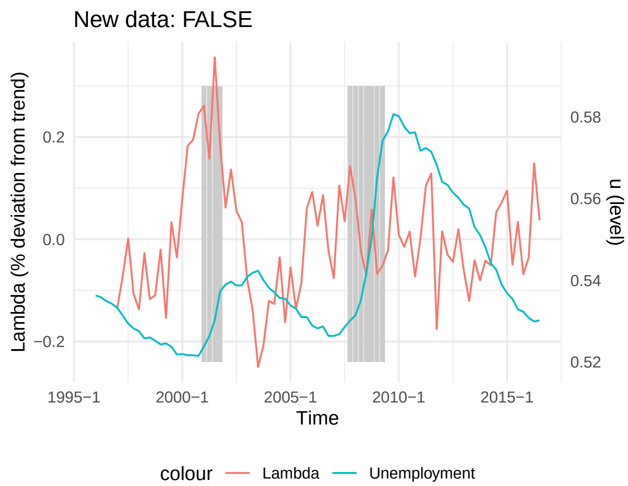


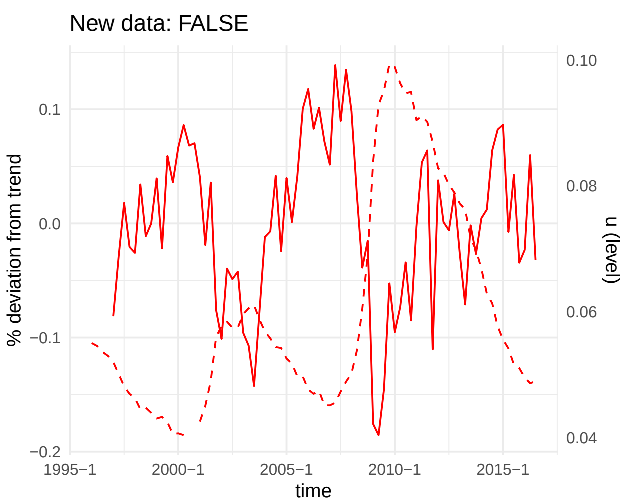
Figure 3: Employed Search Effort Fit

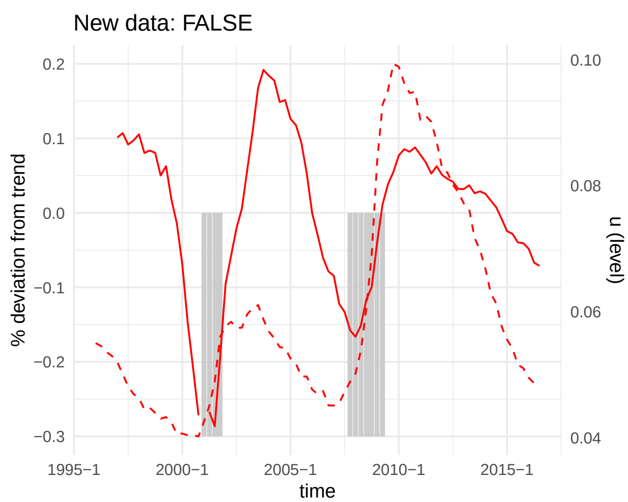
```
##  [1] 0.2000000 0.2315789 0.2631579 0.2947368 0.3263158 0.3578947 0.3894737
##  [8] 0.4210526 0.4526316 0.4842105 0.5157895 0.5473684 0.5789474 0.6105263
## [15] 0.6421053 0.6736842 0.7052632 0.7368421 0.7684211 0.8000000
##  [1] 0.2000000 0.2315789 0.2631579 0.2947368 0.3263158 0.3578947 0.3894737
##  [8] 0.4210526 0.4526316 0.4842105 0.5157895 0.5473684 0.5789474 0.6105263
## [15] 0.6421053 0.6736842 0.7052632 0.7368421 0.7684211 0.8000000
```

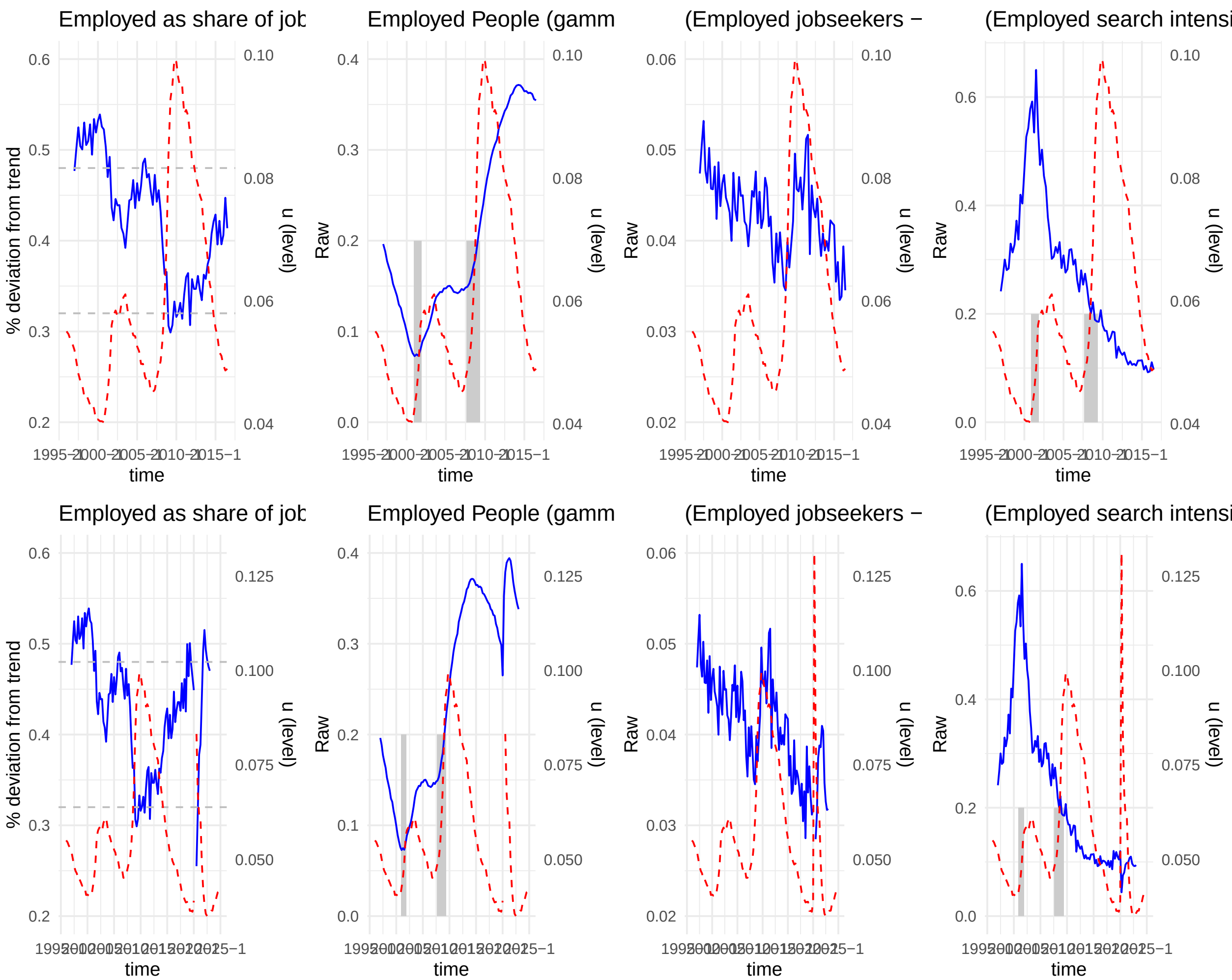


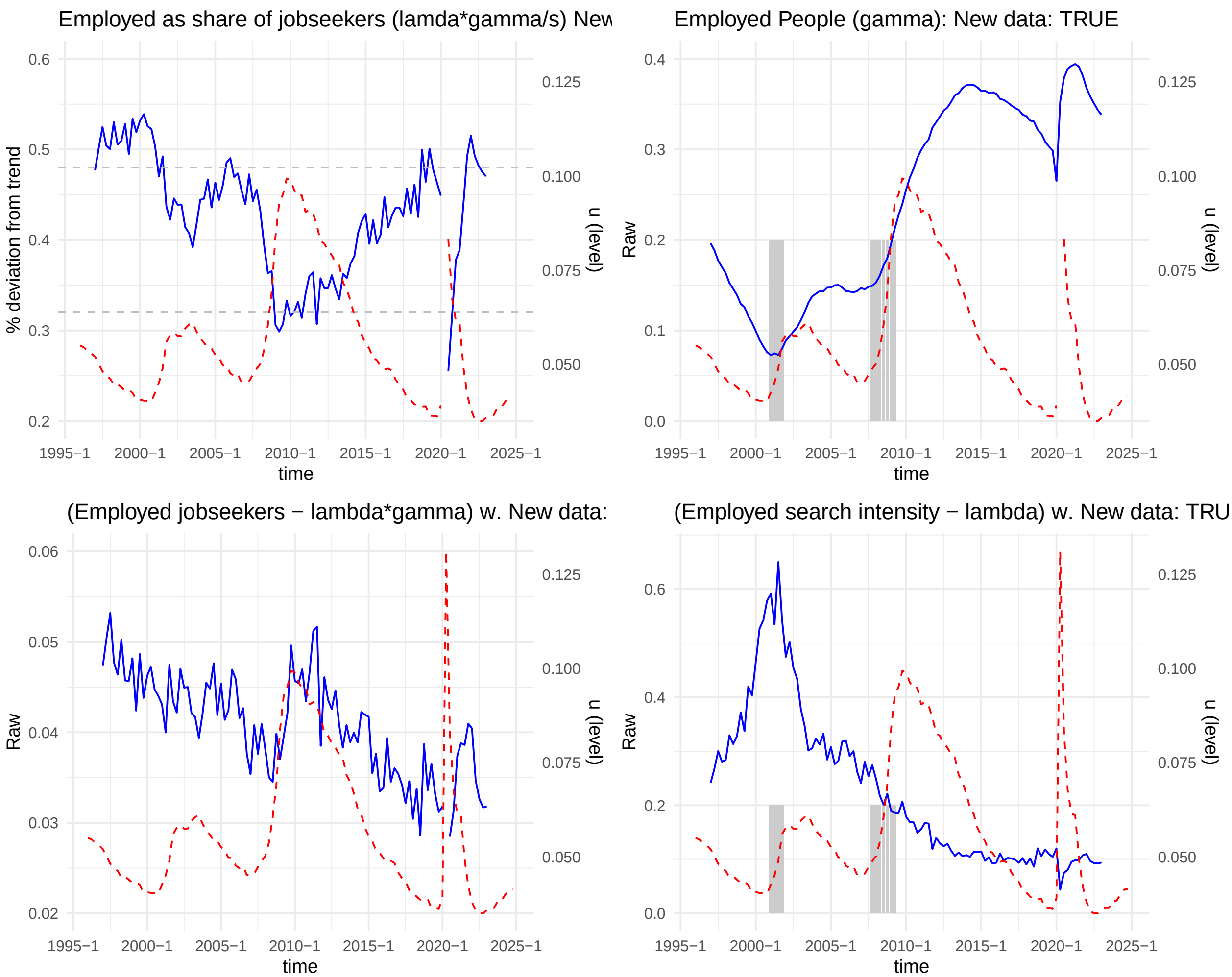

New data: FALSE
Gamma (% deviation from trend)
u (level)
Time
colour Gamma Unemployment


New data: FALSE
Lambda (% deviation from trend)
u (level)
Time
colour Lambda Unemployment


New data: FALSE
% deviation from trend
u (level)
time

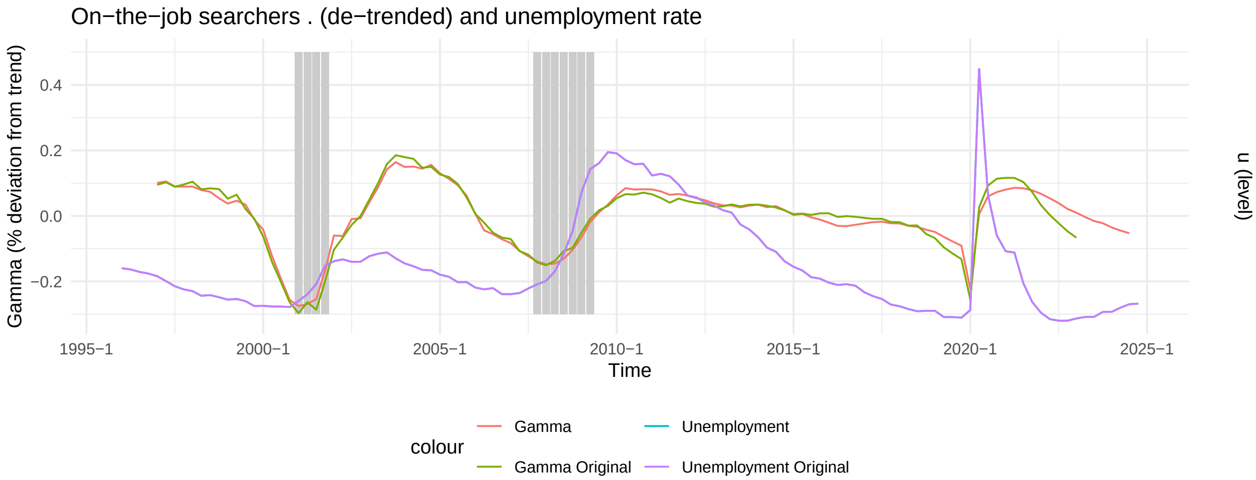

New data: FALSE
% deviation from trend
u (level)
time

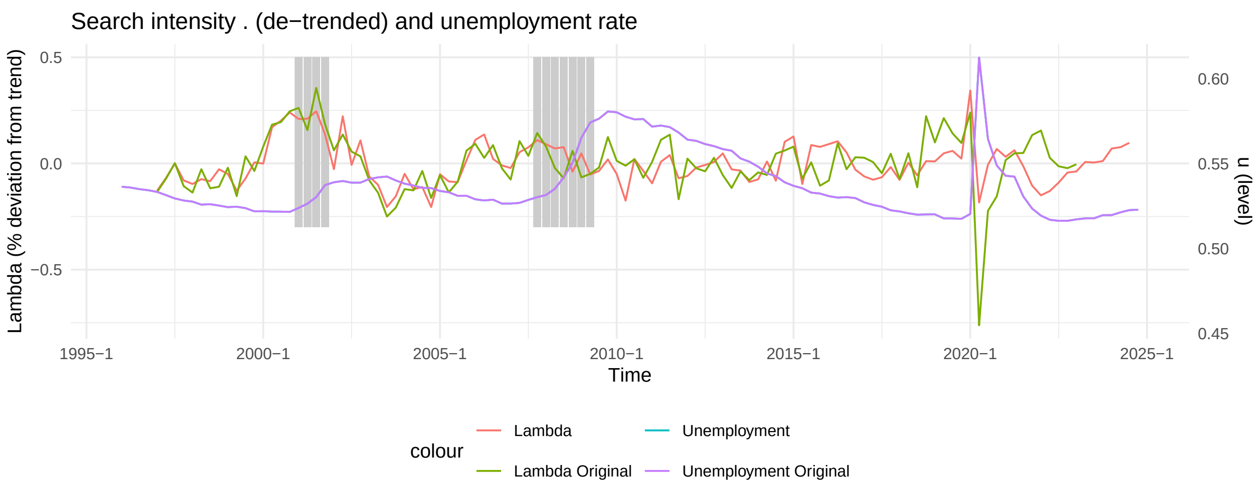

Employed as share of job
Employed People (gamm
(Employed jobseekers −
(Employed search intensi
% deviation from trend
Raw
u (level)
time




```
##
## Call:
## lm(formula = as.formula(forms[which(names(forms) == form)]))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.014882 -0.006066 -0.003639  0.007309  0.026123
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.17447    0.03519  -4.958 3.19e-06 ***
## x            0.23294    0.03499   6.656 1.93e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01036 on 93 degrees of freedom
## Multiple R-squared:  0.3227, Adjusted R-squared:  0.3154
## F-statistic: 44.31 on 1 and 93 DF,  p-value: 1.925e-09
##
##
## Call:
## lm(formula = as.formula(forms[which(names(forms) == form)]))
```

```
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0102059 -0.0031620 -0.0001317  0.0039334  0.0079867
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.268e-02  1.773e-02  -2.970  0.00379 **
## x            1.285e-01  1.731e-02   7.423 5.61e-11 ***
## trend       -3.507e-04  1.918e-05 -18.285  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004839 on 92 degrees of freedom
## Multiple R-squared:  0.8538, Adjusted R-squared:  0.8507
## F-statistic: 268.7 on 2 and 92 DF,  p-value: < 2.2e-16
##
##
## Call:
## lm(formula = as.formula(forms[which(names(forms) == form)]))
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0068610 -0.0016116  0.0001739  0.0018603  0.0046844
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.049270   0.008149  -6.046 3.05e-08 ***
## x            0.049021   0.008104   6.049 3.02e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002399 on 93 degrees of freedom
## Multiple R-squared:  0.2824, Adjusted R-squared:  0.2747
## F-statistic: 36.59 on 1 and 93 DF,  p-value: 3.018e-08
```





Learning Rate

Mueller et al. Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias Mueller et al: Job Seekers' Perceptions and Employment Prospects

The authors claim to disentangle the effects of duration dependence and dynamic selection by using job seekers' elicited beliefs about job-finding. Assuming (and confirming empirically) that job-seekers have realistic initial beliefs about job-finding they isolate the heterogeneity in jobseekers from true duration dependence. Ultimately, they find that dynamic selection selection explains most of the negative duration dependence (rather than pure, true duration dependence).

Findings: Results are remarkably consistent even when including additional data from 2019-2024.

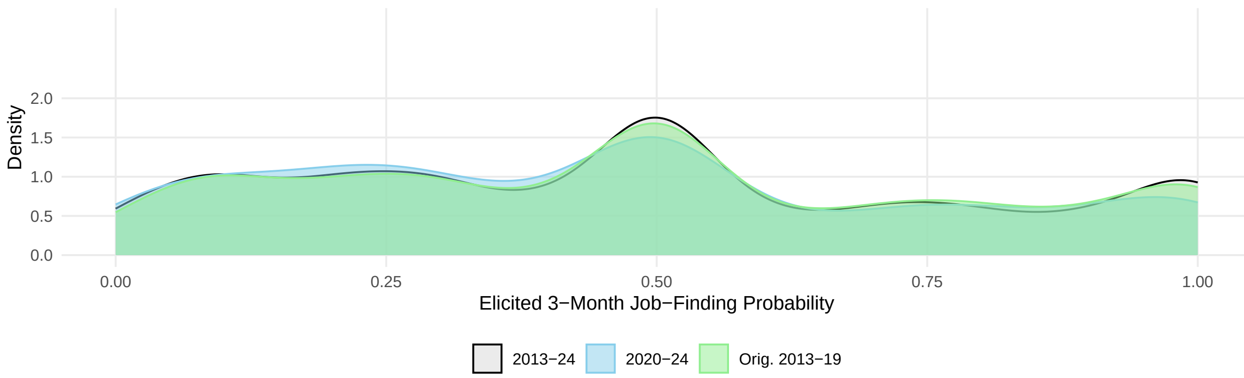
We aim to include this information in our theoretical model of the job search effort as a learning rate (ie. individuals learn about their re-employment probability with repeated failures in the job search).

```
##
## Descriptive Statistics (SCE)
## ============================================================
## Variable                          Orig. 2013-19 2013-24 2020-24
## ------------------------------------------------------------
## High-School Degree or Less             44.5        40.6    36.9
## Some College Education                 32.4        34.9    37.6
```

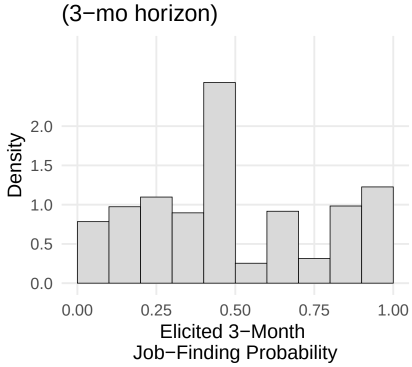
```
## College Degree or More                  23.1       24.6    25.6
## Age 20-34                                25.4       27.2    30.0
## Age 35-49                                33.5       33.6    35.3
## Age 50-65                                41.1       39.2    34.8
## Female                                   59.3       61.2    60.8
## Black                                    19.1       17.9    16.4
## Hispanic                                 12.5       13.0    12.6
## UE transition rate                       18.7       19.1    18.2
## UE transition rate: ST                   25.8       26.5    24.3
## UE transition rate: LT                   12.7       12.7    12.3
## # respondents                             948      1,367    433
## # respondents w/ at least 2 u obs         534       780     252
## # observations                          2,597      3,926   1,347
## ----------------------------------------------------------------
```

Density Comparison of
Elicited Job−Finding Probabilities

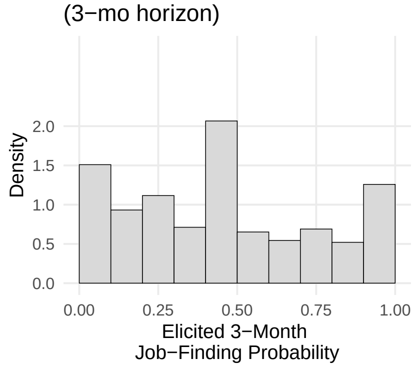
Remarkably consistent beliefs
in job−finding probabilities even when including the Covid period.



Orig. 2013−19: Figure 1.
Histograms of
Elicited Job−Finding
Probabilities −
Panel A. SCE
(3−mo horizon)



2013−24: Figure 1.
Histograms of
Elicited Job−Finding
Probabilities −
Panel A. SCE
(3−mo horizon)



2020−24: Figure 1.
Histograms of
Elicited Job−Finding
Probabilities −
Panel A. SCE
(3−mo horizon)

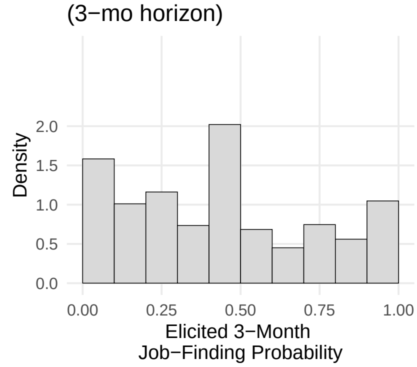
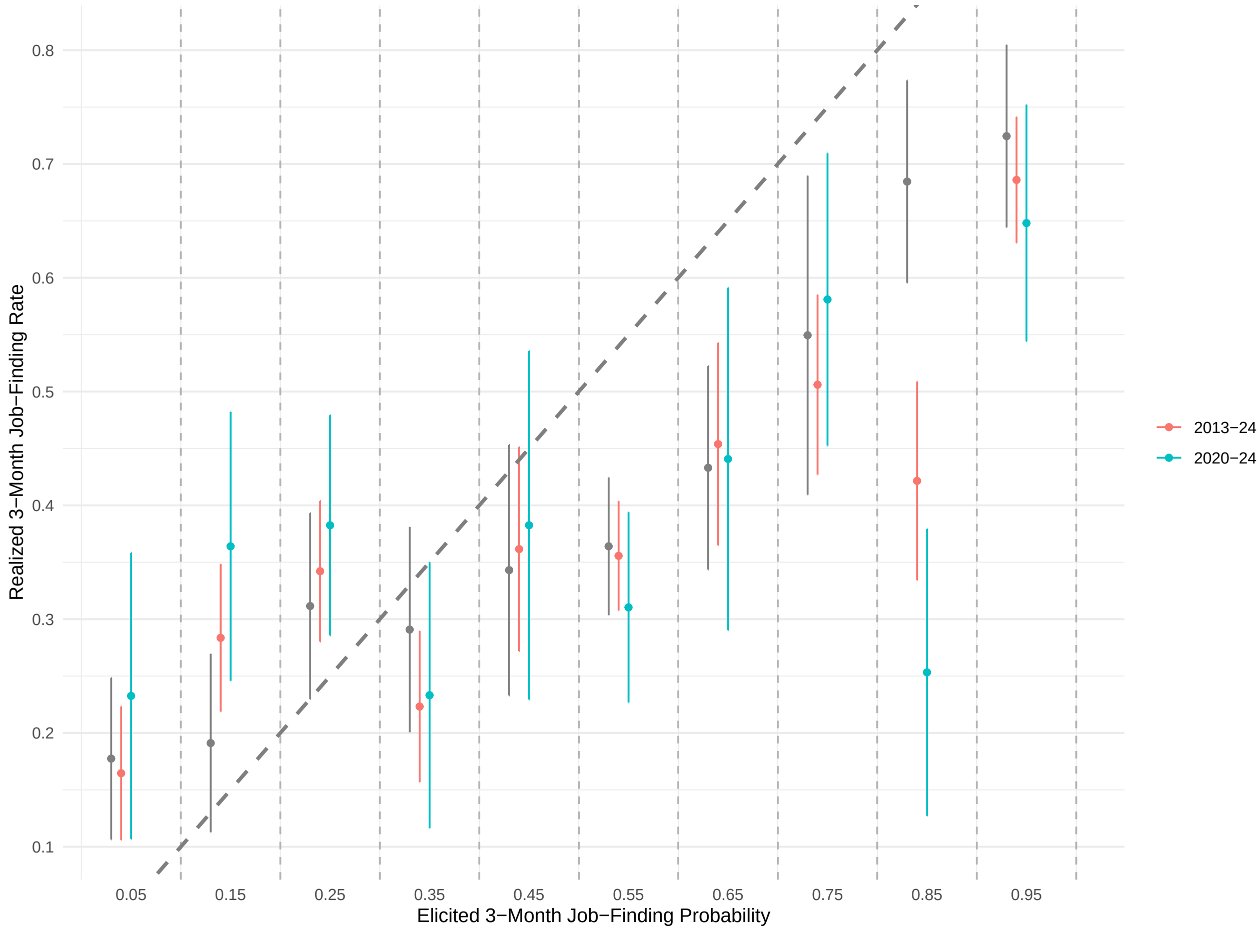


Figure 2: Averages of Realized Job−Finding Rates, by Bins of Elicited Probabilities (SCE)
Demonstrates the predictive power of beliefs. Remarkably consistent even when including the Covid period
except for those who were perhaps overly optimistic.




```
## [1] "Table 2-Regressions of Realized on Elicited 3-Month Job-Finding Probabilities (SCE)"
## [1] "Panel A. Contemporaneous elicitations"

##
## ========================================================================
##                                       Dependent variable:
##                          -----------------------------------------------
##                                   T+3 UE Transitions (3-Months)
##                            Orig. 2013-19        2013-24         2020-24
##                                 (1)               (2)             (3)
## ------------------------------------------------------------------------
## find_job_3mon               0.464***          0.396***        0.265***
##                              (0.045)           (0.036)         (0.067)
##
## 1 | userid
##
##
## Constant                     -0.104            -0.080          -0.136
##                              (0.169)           (0.137)         (0.267)
##
## ------------------------------------------------------------------------
```

```
## Observations              1,201            1,911            673
## R2                        0.218            0.139            0.105
## Adjusted R2               0.207            0.132            0.083
## Residual Std. Error 0.467 (df = 1184) 0.475 (df = 1894) 0.478 (df = 656)
## ===========================================================================
## Note:                                          *p<0.1; **p<0.05; ***p<0.01

##
## =============================================================================
##                                          Dependent variable:
##                       -------------------------------------------------------
##                                     T+3 UE Transitions (3-Months)
##                         Orig. 2013-19        2013-24          2020-24
##                              (1)               (2)              (3)
## -----------------------------------------------------------------------------
## find_job_3mon              0.501***          0.418***         0.391***
##                            (0.061)           (0.051)          (0.094)
##
## findjob_3mon_longterm     -0.258***         -0.170**         -0.360***
##                            (0.088)           (0.071)          (0.133)
##
## longterm_unemployed        -0.078           -0.127***         -0.043
##                            (0.051)           (0.041)          (0.075)
##
## 1 | userid
##
##
## Constant                   -0.062            -0.063           -0.402
##                            (0.175)           (0.139)          (0.266)
##
## -----------------------------------------------------------------------------
## Observations                1,201             1,911            673
## R2                          0.259             0.182            0.155
## Adjusted R2                 0.248             0.174            0.132
## Residual Std. Error   0.455 (df = 1182) 0.464 (df = 1892) 0.465 (df = 654)
## =============================================================================
## Note:                                          *p<0.1; **p<0.05; ***p<0.01
## [1] "Panel B. Lagged elicitations"

##
## =========================================================================
##                                        Dependent variable:
##                     -----------------------------------------------------
##                                   T+3 UE Transitions (3-Months)
##                       Orig. 2013-19       2013-24          2020-24
##                            (1)              (2)              (3)
## -------------------------------------------------------------------------
## tplus3_percep_3mon      0.332***         0.241***         0.203**
##                         (0.067)          (0.056)          (0.102)
##
## 1 | userid
##
##
## Constant                 0.304           0.490**           0.451
```

```
##                        (0.270)          (0.207)          (0.394)
##
## ---------------------------------------------------------------------
## Observations             474              798              300
## R2                      0.168            0.090            0.179
## Adjusted R2             0.139            0.071            0.132
## Residual Std. Error 0.398 (df = 457) 0.436 (df = 781) 0.447 (df = 283)
## =====================================================================
## Note:                                     *p<0.1; **p<0.05; ***p<0.01
```

```
##
## =====================================================================
##                                Dependent variable:
##                     -------------------------------------------------
##                              T+3 UE Transitions (3-Months)
##                      Orig. 2013-19       2013-24          2020-24
##                           (1)              (2)              (3)
## ---------------------------------------------------------------------
## find_job_3mon          0.301***         0.205***          -0.035
##                        (0.069)          (0.058)          (0.110)
##
## 1 | userid
##
##
## Constant                0.201           0.422**           0.361
##                        (0.274)          (0.207)          (0.400)
##
## ---------------------------------------------------------------------
## Observations             474              798              300
## R2                      0.159            0.083            0.168
## Adjusted R2             0.129            0.064            0.121
## Residual Std. Error 0.400 (df = 457) 0.437 (df = 781) 0.450 (df = 283)
## =====================================================================
## Note:                                     *p<0.1; **p<0.05; ***p<0.01
```

Fig 3. Perceived vs. Realized Job Finding, by Duration of Unemployment

Duration dependence is strongly negative across all samples.
Bias in beliefs of LTUE is also consistently high across samples.

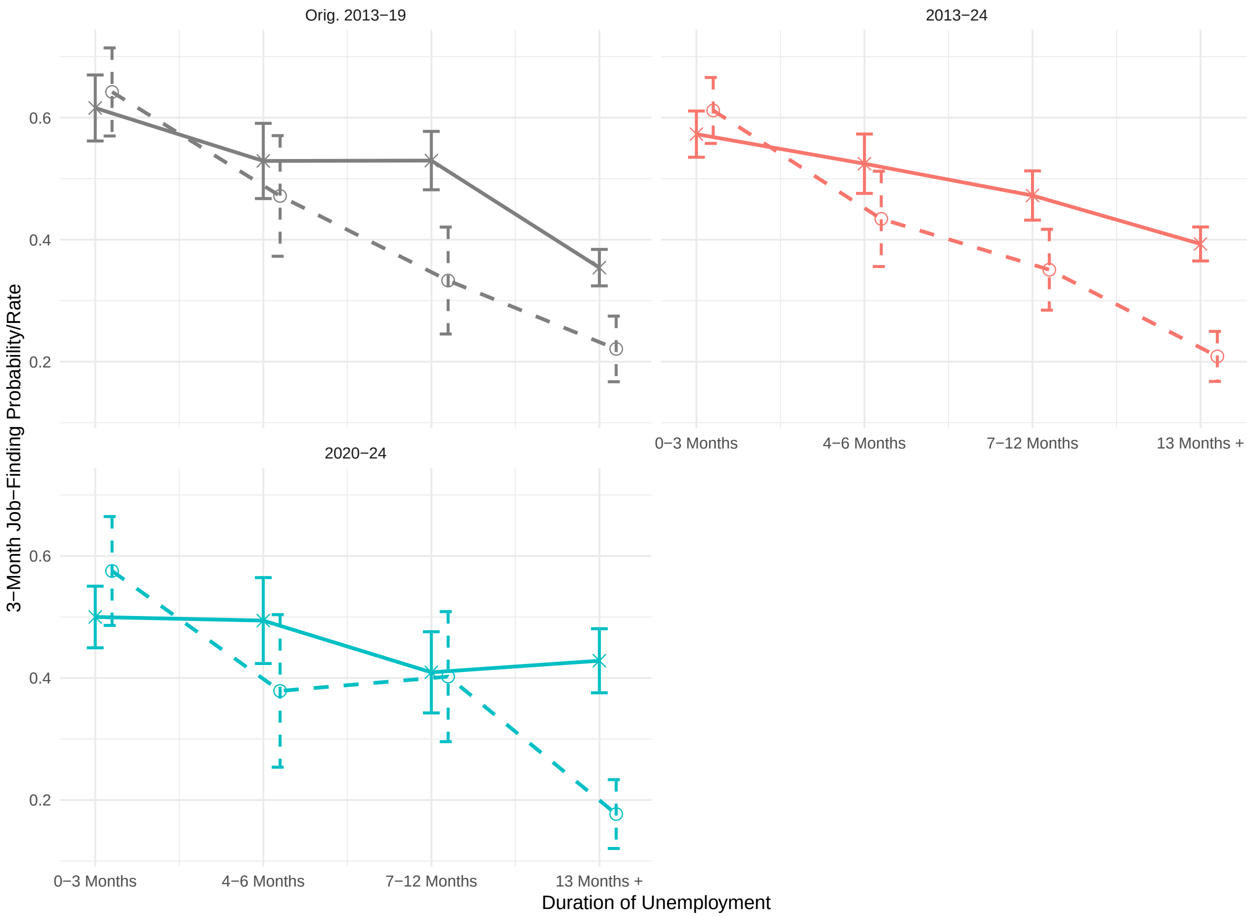
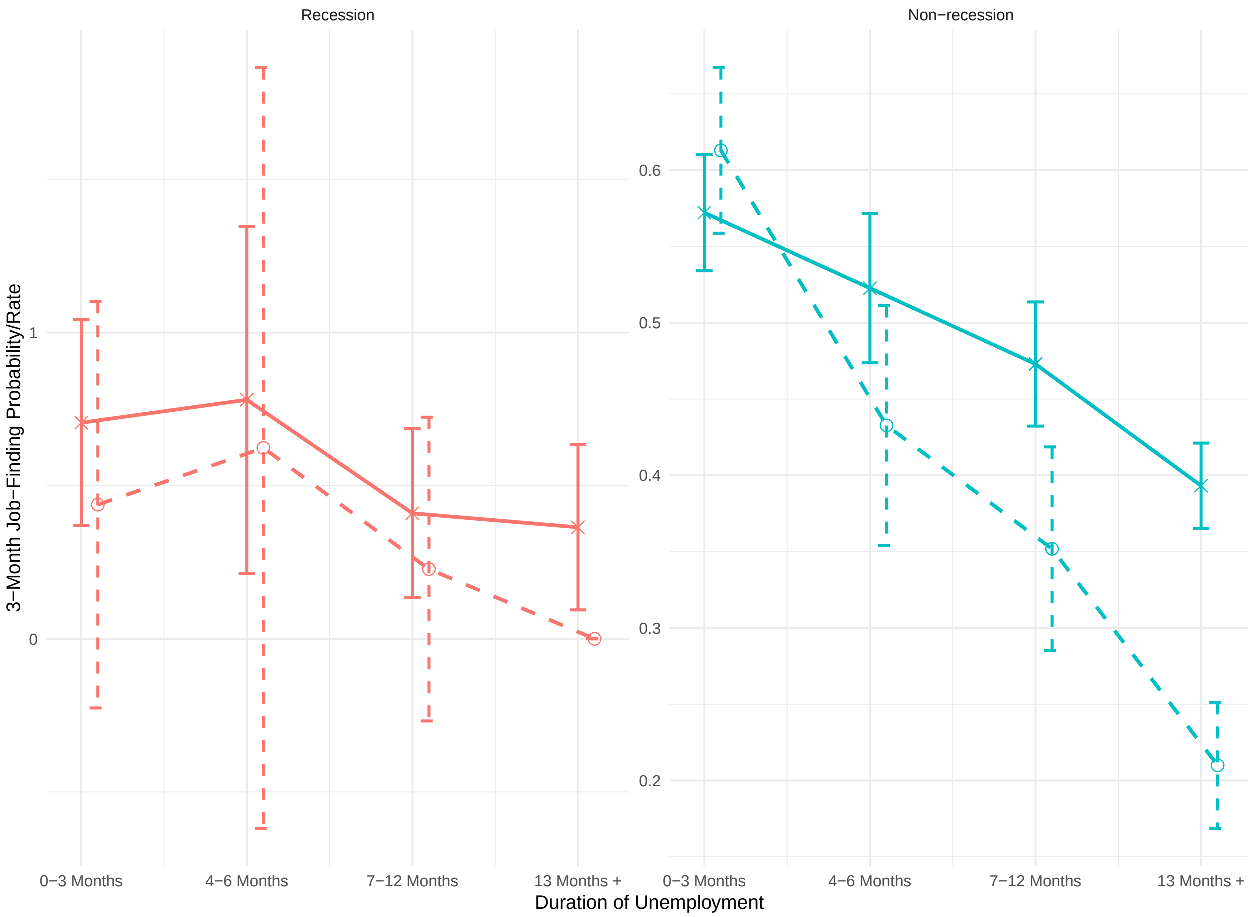


Fig 3. Perceived vs. Realized Job Finding, by Duration of Unemployment

Duration dependence is strongly negative across all samples.
Bias in beliefs of LTUE is also consistently high across samples.




```
## [1] "Table 4-Linear Regressions of Elicited Job-Finding Probabilities on Duration of Unemployment"
```

	(1)	(2)	(3)	(4)
Orig. 2013-19				
Unemployment Duration (Months)	-0.0057	-0.0050	-0.0043	0.0022
	(0.0007)	(0.0007)	(0.0006)	(0.0049)
Num.Obs.	882	2281	2281	2281
R2	0.110	0.090	0.155	0.824
2013-24				
Unemployment Duration (Months)	-0.0050	-0.0048	-0.0042	-0.0026
	(0.0006)	(0.0006)	(0.0005)	(0.0034)
Num.Obs.	1265	3423	3399	3423

```
## | R2                               | 0.067    | 0.065    | 0.109    | 0.817    |
## +----------------------------------+----------+----------+----------+----------+
## | 2020-24                                                                      |
## +----------------------------------+----------+----------+----------+----------+
## | Unemployment Duration (Months)   | -0.0011  | -0.0035  | -0.0039  | -0.0077  |
## +----------------------------------+----------+----------+----------+----------+
## |                                  | (0.0013) | (0.0012) | (0.0013) | (0.0036) |
## +----------------------------------+----------+----------+----------+----------+
## | Num.Obs.                         | 395      | 1150     | 1140     | 1150     |
## +----------------------------------+----------+----------+----------+----------+
## | R2                               | 0.002    | 0.019    | 0.118    | 0.838    |
## +==================================+==========+==========+==========+==========+
## | Standard errors are clustered at the user or spell level as indicated.       |
## +==================================+==========+==========+==========+==========+
## Table: Table 4 - Panel A: Linear Regressions of Elicited Job-Finding Probabilities on Duration of Une
```

Fig 4. Changes in Job−Finding Probability Across and Within Spells

Figure 4 illustrates the difference between the observed (cross−sectional − left panel) duration dependence and the true (individual−level − right panel) duration dependence in the reported beliefs graphically.

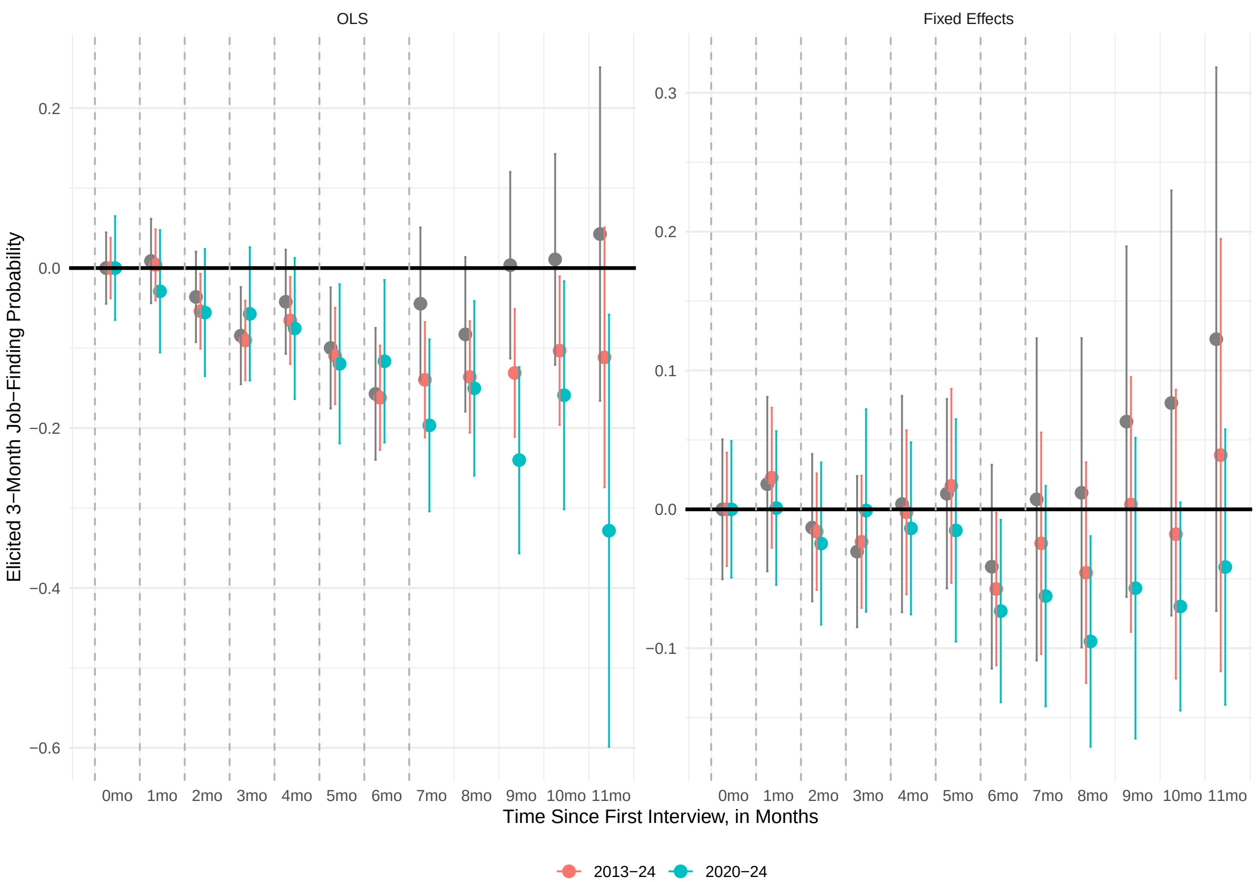


Fig 3. Perceived vs. Realized Job Finding, by Duration of Unemployment

Duration dependence is strongly negative across all samples.
Bias in beliefs of LTUE is also consistently high across samples.

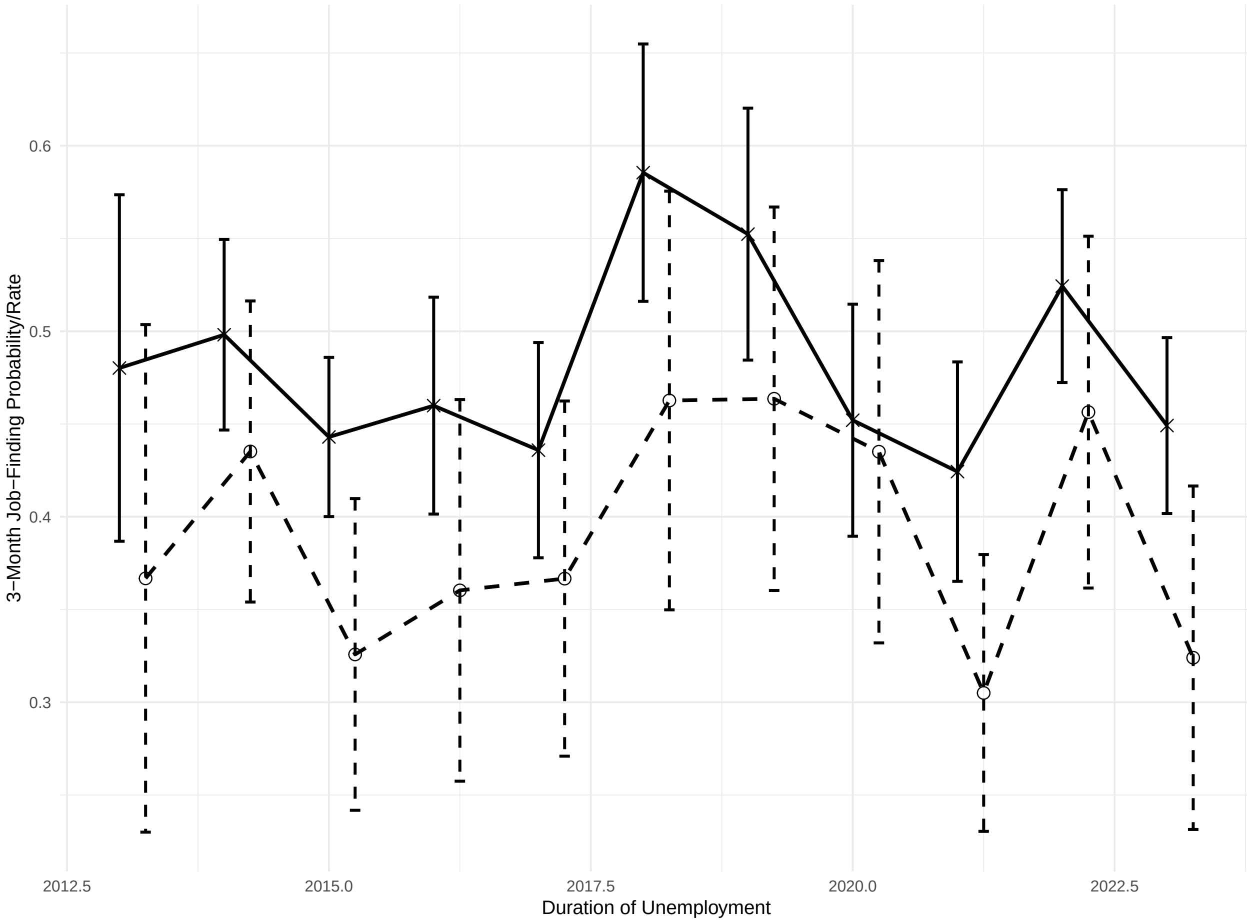


Fig 3. Perceived and Realized Job Finding, by Year

Discarded Analyses

1. (Replicated with additional data for unemployed jobseekers) **Mukoyama et al. 2018: Job Search Over the Business Cycle**

They provide a novel measure of job search effort exploiting the American Time Use and Current Population Surveys which can be reduced to just the intensive margin (changes in search effort by worker!). At the moment, I think this will be the most useful input for our model.

Abstract: We examine the cyclicality of search effort using time-series, cross-state, and individual variation and find that it is countercyclical. We then set up a search and matching model with endogenous search effort and show that search effort does not amplify labor market fluctuations but rather dampens them. Lastly, we examine the role of search effort in driving recent unemployment dynamics and show that the unemployment rate would have been 0.5 to 1 percentage points higher in the 2008–2014 period had search effort not increased.

4. **Survey of Consumer Expectations Reservation Wages, Accepted Wages, and Wage Expectations** The data is unfortunately sparse and linking outcomes to reservation wages is difficult. However, in a cross-sectional setting we are able to deduce some weak relationships between Unemployment Duration and Absolute Reservation Wages and Wage Expectations.*

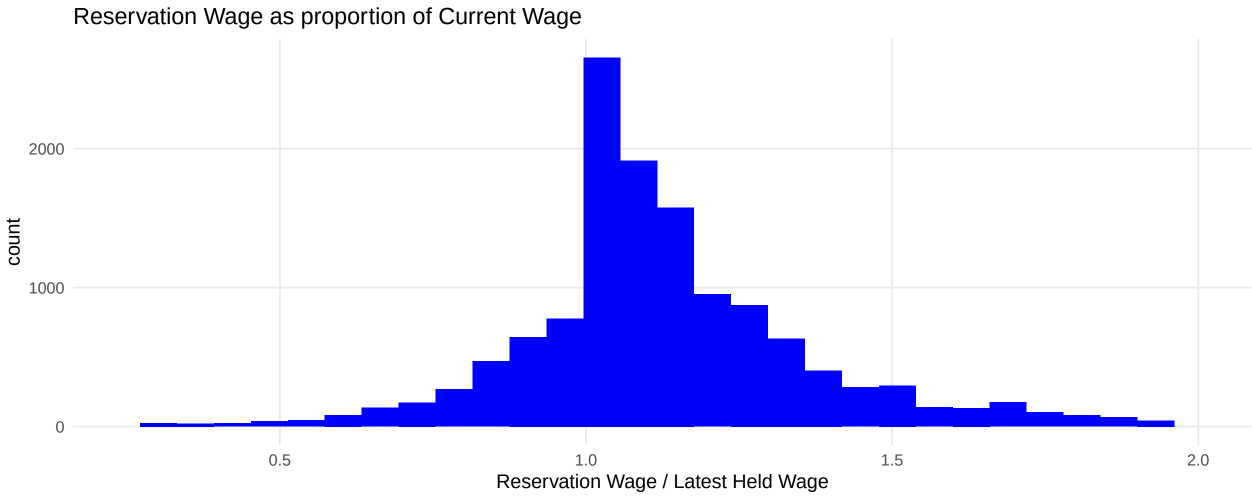
Reservation Wages

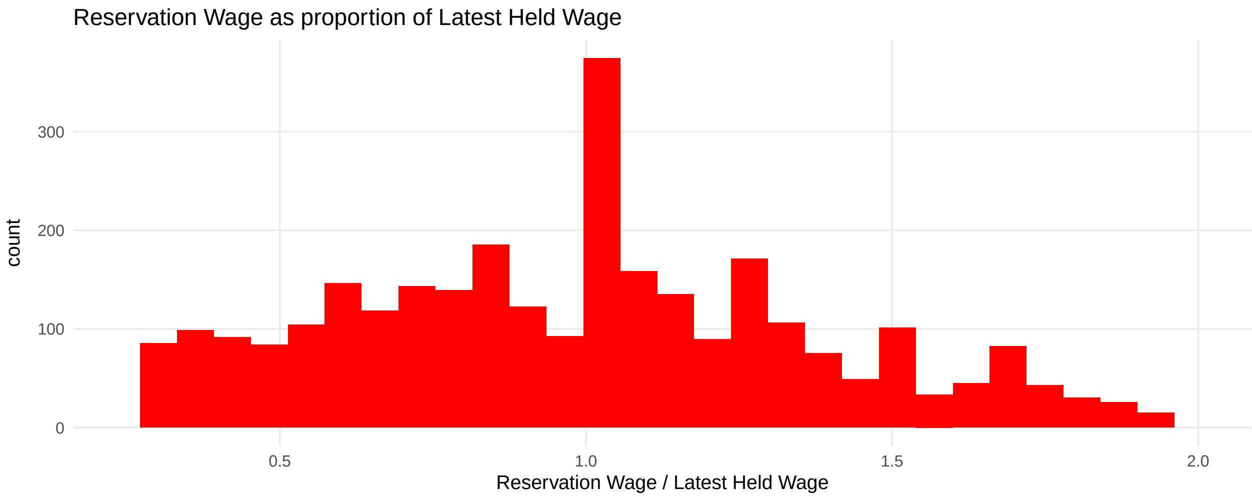
Exploring the effect of unemployment duration on reservation wages, accepted wages, and expected wage offers.

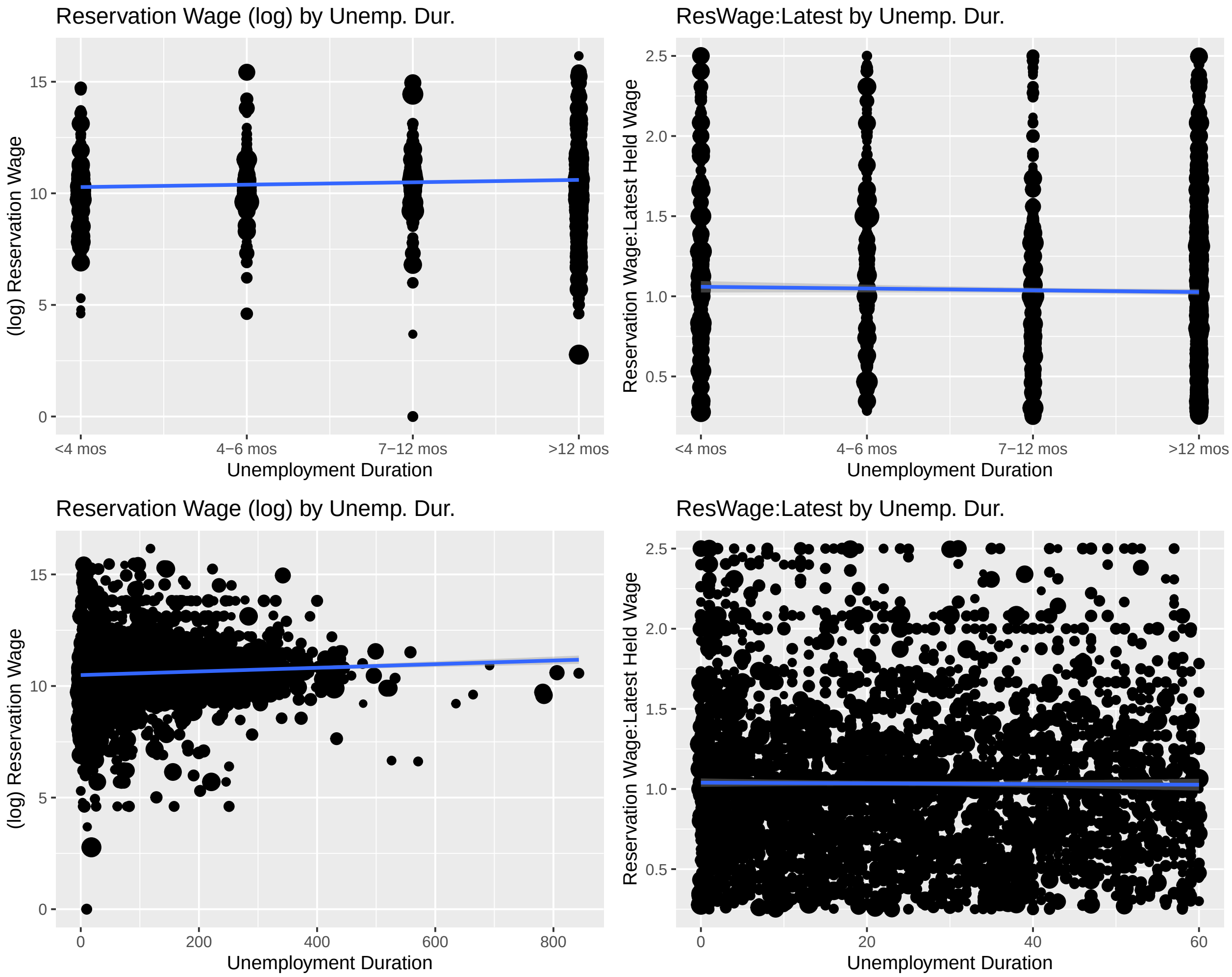
Survey of Consumer Expectations Reservation Wages, Accepted Wages, and Wage Expectations (2014-2022) *The data is unfortunately sparse and linking outcomes to reservation wages is difficult. However, in a cross-sectional setting we are able to deduce some weak relationships between Unemployment Duration and Absolute Reservation Wages and Wage Expectations.*

```
source(here("data/behav_params/SCE Labour Market Survey/sce_res_wage_analysis.R"))
```

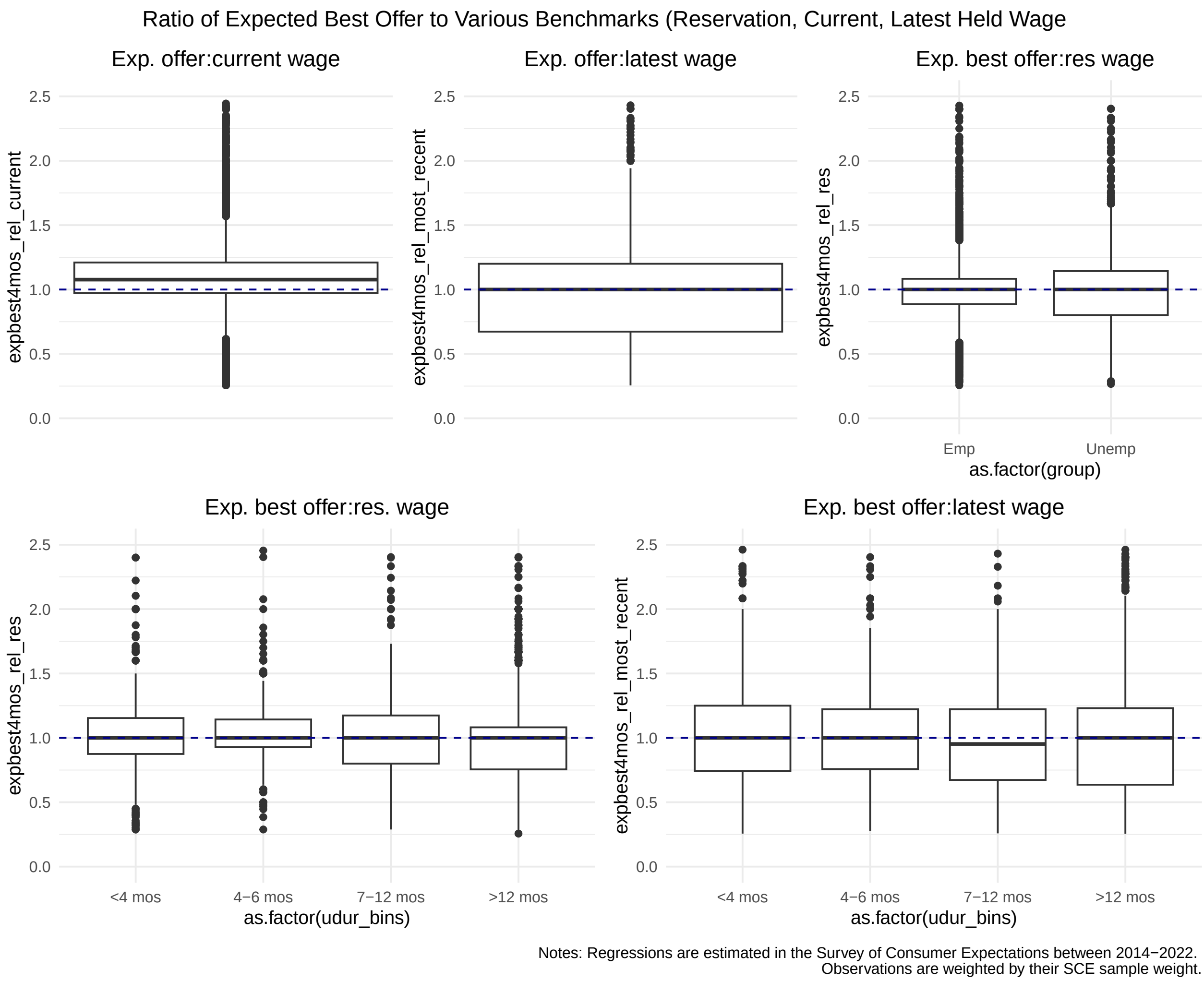
```
## [1] "Plots of RESERVATION WAGE versus latest, current wage"
```



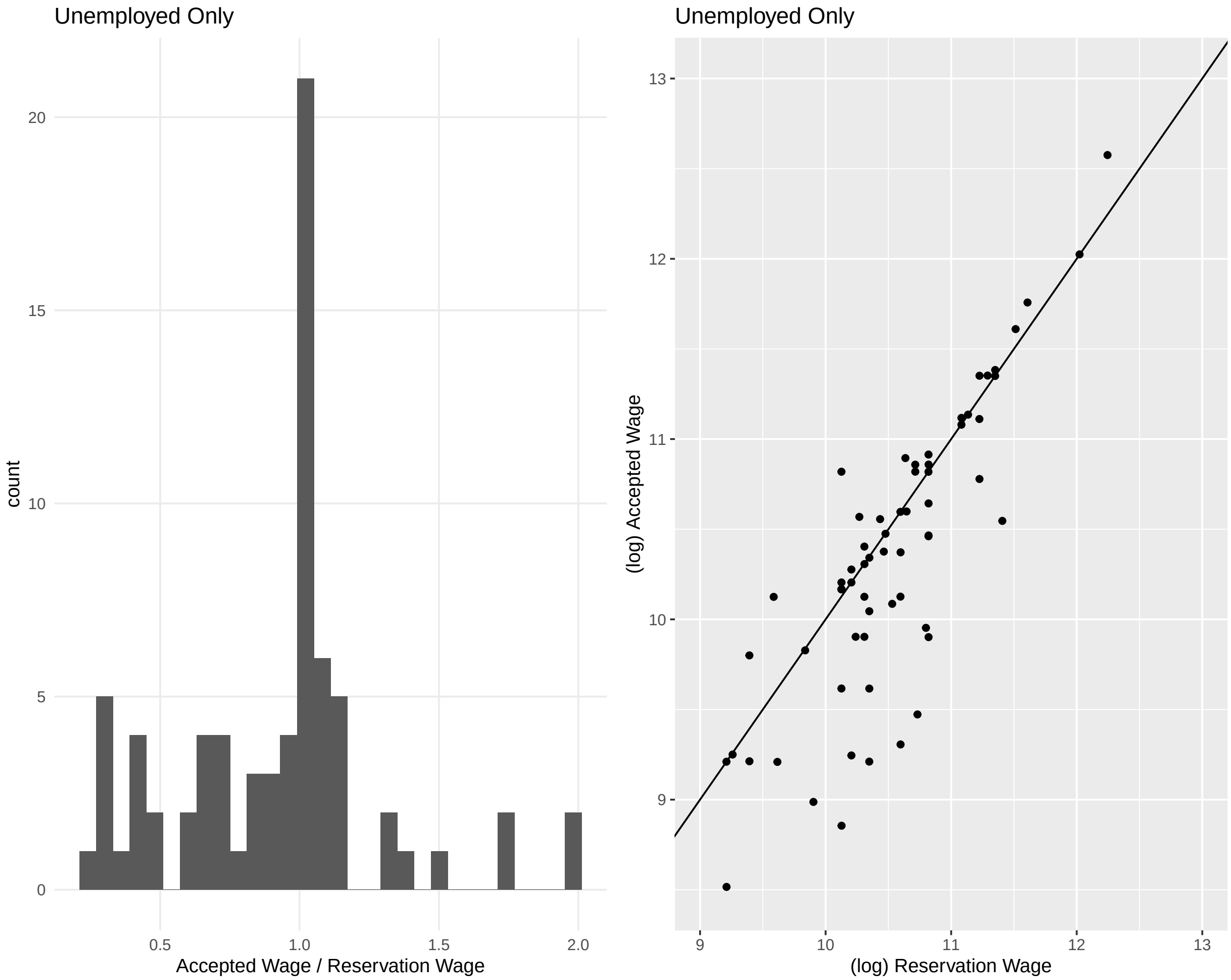


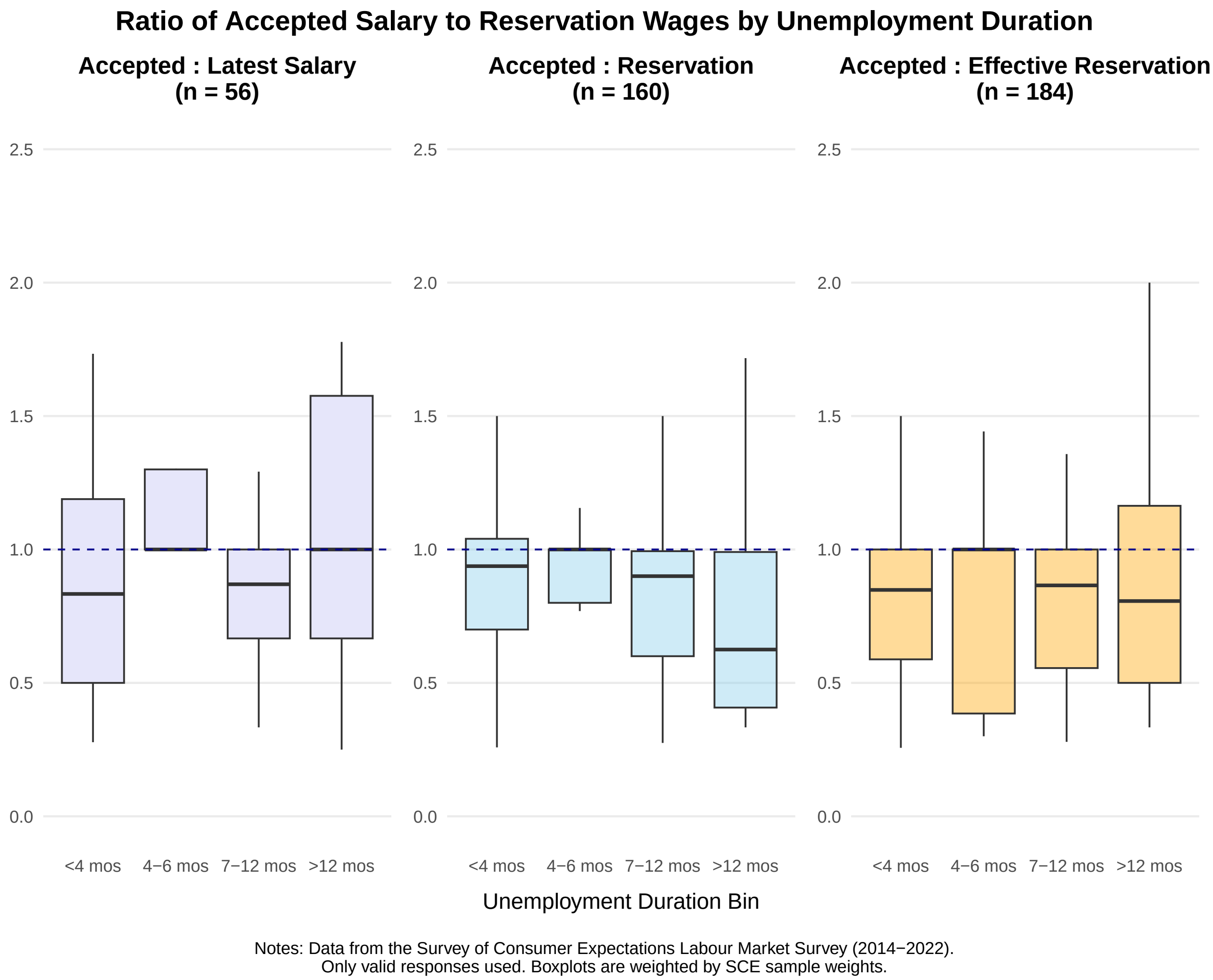



```
## [1] "Plots of EXPECTED OFFER versus latest, current, reservation wage"
```




```
## [1] "Plots of ACCEPTED SALARY versus latest, current, reservation wage"
```


Unemployed Only
count
20
15
10
5
0
0.5
1.0
1.5
2.0
Accepted Wage / Reservation Wage
Unemployed Only
(log) Accepted Wage
13
12
11
10
9
9
10
11
12
13
(log) Reservation Wage



	Accpt:Latest	AccptWage w.c	Accpt:ResWage	AccptWage:ResWage w.c	Accpt:EffResW
(Intercept)	0.826***	1.743***	0.933***	1.199***	0.826***
	(0.108)	(0.260)	(0.045)	(0.141)	(0.051)
udur_bins	0.050	-0.005	-0.048*	-0.053**	0.008
	(0.045)	(0.048)	(0.019)	(0.020)	(0.023)
Num.Obs.	56	56	160	159	184
R2	0.022	0.430	0.040	0.118	0.001
RMSE	0.40	0.35	0.30	0.30	0.34

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table: Accepted Wages and Unemployment Duration

```
## \begin{table}
## \centering
## \begin{talltblr}[         %% tabularray outer open
## caption={Reservation Wages and Unemployment Duration},
## note{}={+ p \num{< 0.1}, * p \num{< 0.05}, ** p \num{< 0.01}, *** p \num{< 0.001}},
## ]                     %% tabularray outer close
## {                     %% tabularray inner open
## colspec={Q[]Q[]Q[]Q[]Q[]},
## column{2,3,4,5}={}{halign=c,},
## column{1}={}{halign=l,},
## hline{26}={1,2,3,4,5}{solid, black, 0.05em},
## }                     %% tabularray inner close
## \toprule
## & ResWage & ResWage w.c & ResWage/LastWage & ResWage/LastWage w.c \\ \midrule %% TinyTableHeader
## (Intercept) & \num{10.173}*** & \num{9.945}*** & \num{0.825}*** & \num{0.750}*** \\
## & (\num{0.044}) & (\num{0.071}) & (\num{0.022}) & (\num{0.040}) \\
## udur\_bins & \num{0.107}*** & \num{0.083}*** & \num{0.027}*** & \num{0.023}*** \\
## & (\num{0.012}) & (\num{0.011}) & (\num{0.006}) & (\num{0.006}) \\
## female &  & \num{-0.275}*** &  & \num{0.009} \\
## &  & (\num{0.022}) &  & (\num{0.012}) \\
## age &  & \num{0.005}*** &  & \num{0.001}* \\
## &  & (\num{0.001}) &  & (\num{0.000}) \\
## hhinc\_2 &  & \num{0.230}*** &  & \num{-0.008} \\
## &  & (\num{0.026}) &  & (\num{0.014}) \\
## hhinc\_3 &  & \num{0.427}*** &  & \num{-0.017} \\
## &  & (\num{0.030}) &  & (\num{0.017}) \\
## hhinc\_4 &  & \num{0.759}*** &  & \num{-0.008} \\
## &  & (\num{0.033}) &  & (\num{0.019}) \\
## education\_2 &  & \num{-0.247}*** &  & \num{0.050}+ \\
## &  & (\num{0.045}) &  & (\num{0.026}) \\
## education\_3 &  & \num{-0.122}** &  & \num{0.007} \\
## &  & (\num{0.047}) &  & (\num{0.027}) \\
## education\_4 &  & \num{-0.046} &  & \num{0.052}+ \\
## &  & (\num{0.051}) &  & (\num{0.029}) \\
## education\_5 &  & \num{0.027} &  & \num{0.008} \\
## &  & (\num{0.049}) &  & (\num{0.028}) \\
## education\_6 &  & \num{0.111}* &  & \num{0.054}+ \\
## &  & (\num{0.054}) &  & (\num{0.031}) \\
## Num.Obs. & \num{7937} & \num{7824} & \num{6294} & \num{6224} \\
## R2 & \num{0.010} & \num{0.169} & \num{0.003} & \num{0.007} \\
## R2 Adj. & \num{0.010} & \num{0.168} & \num{0.003} & \num{0.005} \\
## AIC & \num{191435.4} & \num{187281.4} & \num{9054.4} & \num{8961.7} \\
## BIC & \num{191456.4} & \num{187372.0} & \num{9074.6} & \num{9049.3} \\
## Log.Lik. & \num{-11923.451} & \num{-11075.843} & \num{-4524.195} & \num{-4467.857} \\
## RMSE & \num{0.98} & \num{0.90} & \num{0.44} & \num{0.44} \\
## \bottomrule
## \end{talltblr}
## \end{table}
## \begin{table}
## \centering
## \begin{talltblr}[         %% tabularray outer open
## caption={Accepted Wages and Unemployment Duration},
## note{}={+ p \num{< 0.1}, * p \num{< 0.05}, ** p \num{< 0.01}, *** p \num{< 0.001}},
## ]                     %% tabularray outer close
```

```
## {                      %% tabularray inner open
## colspec={Q[]Q[]Q[]Q[]Q[]},
## column{2,3,4,5}={}{halign=c,},
## column{1}={}{halign=l,},
## hline{26}={1,2,3,4,5}{solid, black, 0.05em},
## }                      %% tabularray inner close
## \toprule
## & AccptWage & AccptWage w.c & AccptWage/ResWage & AccptWage/ResWage w.c \\ \midrule %% TinyTableHeade
## (Intercept) & \num{10.568}*** & \num{11.705}*** & \num{0.924}*** & \num{1.303}*** \\
## & (\num{0.106}) & (\num{0.255}) & (\num{0.048}) & (\num{0.132}) \\
## udur\_bins & \num{-0.006} & \num{-0.037} & \num{-0.031}+ & \num{-0.036}+ \\
## & (\num{0.040}) & (\num{0.037}) & (\num{0.018}) & (\num{0.018}) \\
## female &  & \num{-0.164} &  & \num{-0.073} \\
## &  & (\num{0.102}) &  & (\num{0.050}) \\
## age &  & \num{-0.008}* &  & \num{-0.005}** \\
## &  & (\num{0.004}) &  & (\num{0.002}) \\
## hhinc\_2 &  & \num{0.260}+ &  & \num{0.043} \\
## &  & (\num{0.139}) &  & (\num{0.067}) \\
## hhinc\_3 &  & \num{0.272}+ &  & \num{0.042} \\
## &  & (\num{0.138}) &  & (\num{0.069}) \\
## hhinc\_4 &  & \num{0.377}* &  & \num{-0.052} \\
## &  & (\num{0.150}) &  & (\num{0.075}) \\
## education\_2 &  & \num{-0.996}*** &  & \num{-0.043} \\
## &  & (\num{0.224}) &  & (\num{0.122}) \\
## education\_3 &  & \num{-0.940}*** &  & \num{-0.128} \\
## &  & (\num{0.223}) &  & (\num{0.122}) \\
## education\_4 &  & \num{-1.036}*** &  & \num{-0.176} \\
## &  & (\num{0.226}) &  & (\num{0.123}) \\
## education\_5 &  & \num{-0.827}*** &  & \num{-0.141} \\
## &  & (\num{0.224}) &  & (\num{0.124}) \\
## education\_6 &  & \num{-0.551}* &  & \num{-0.095} \\
## &  & (\num{0.228}) &  & (\num{0.127}) \\
## Num.Obs. & \num{127} & \num{126} & \num{164} & \num{163} \\
## R2 & \num{0.000} & \num{0.299} & \num{0.017} & \num{0.133} \\
## R2 Adj. & \num{-0.008} & \num{0.232} & \num{0.011} & \num{0.070} \\
## AIC & \num{2933.2} & \num{2884.9} & \num{110.6} & \num{109.9} \\
## BIC & \num{2941.7} & \num{2921.7} & \num{119.9} & \num{150.1} \\
## Log.Lik. & \num{-123.204} & \num{-99.911} & \num{-52.283} & \num{-41.957} \\
## RMSE & \num{0.58} & \num{0.53} & \num{0.32} & \num{0.32} \\
## \bottomrule
## \end{talltblr}
## \end{table}
## \begin{table}
## \centering
## \begin{talltblr}[         %% tabularray outer open
## caption={Expected Wages and Unemployment Duration},
## note{}={+ p \num{< 0.1}, * p \num{< 0.05}, ** p \num{< 0.01}, *** p \num{< 0.001}},
## ]                      %% tabularray outer close
## {                      %% tabularray inner open
## colspec={Q[]Q[]Q[]Q[]Q[]},
## column{2,3,4,5}={}{halign=c,},
## column{1}={}{halign=l,},
## hline{26}={1,2,3,4,5}{solid, black, 0.05em},
## }                      %% tabularray inner close
```

```
## \toprule
## & ExpWage/ResWage & ExpWage/ResWage w.c & ExpWage/LastWage & ExpWage/LastWage w.c \\ \midrule %% Tiny
## (Intercept) & \num{1.057}*** & \num{1.226}*** & \num{1.087}*** & \num{1.257}*** \\
## & (\num{0.020}) & (\num{0.040}) & (\num{0.029}) & (\num{0.059}) \\
## udur\_bins & \num{-0.022}*** & \num{-0.009} & \num{-0.024}** & \num{-0.008} \\
## & (\num{0.006}) & (\num{0.006}) & (\num{0.008}) & (\num{0.009}) \\
## female &  & \num{-0.022}+ &  & \num{0.064}*** \\
## &  & (\num{0.013}) &  & (\num{0.019}) \\
## age &  & \num{-0.003}*** &  & \num{-0.004}*** \\
## &  & (\num{0.000}) &  & (\num{0.001}) \\
## hhinc\_2 &  & \num{0.004} &  & \num{-0.038} \\
## &  & (\num{0.016}) &  & (\num{0.023}) \\
## hhinc\_3 &  & \num{0.004} &  & \num{-0.001} \\
## &  & (\num{0.018}) &  & (\num{0.026}) \\
## hhinc\_4 &  & \num{0.000} &  & \num{-0.005} \\
## &  & (\num{0.019}) &  & (\num{0.027}) \\
## education\_2 &  & \num{-0.035} &  & \num{-0.032} \\
## &  & (\num{0.027}) &  & (\num{0.040}) \\
## education\_3 &  & \num{-0.008} &  & \num{-0.056} \\
## &  & (\num{0.028}) &  & (\num{0.041}) \\
## education\_4 &  & \num{0.004} &  & \num{-0.031} \\
## &  & (\num{0.030}) &  & (\num{0.044}) \\
## education\_5 &  & \num{0.011} &  & \num{-0.090}* \\
## &  & (\num{0.029}) &  & (\num{0.042}) \\
## education\_6 &  & \num{0.021} &  & \num{0.002} \\
## &  & (\num{0.032}) &  & (\num{0.046}) \\
## Num.Obs. & \num{3114} & \num{3070} & \num{2721} & \num{2690} \\
## R2 & \num{0.005} & \num{0.028} & \num{0.003} & \num{0.029} \\
## R2 Adj. & \num{0.005} & \num{0.024} & \num{0.003} & \num{0.025} \\
## AIC & \num{2803.9} & \num{2733.2} & \num{4079.4} & \num{3986.5} \\
## BIC & \num{2822.1} & \num{2811.6} & \num{4097.2} & \num{4063.1} \\
## Log.Lik. & \num{-1398.968} & \num{-1353.588} & \num{-2036.722} & \num{-1980.241} \\
## RMSE & \num{0.34} & \num{0.34} & \num{0.46} & \num{0.45} \\
## \bottomrule
## \end{talltblr}
## \end{table}
```

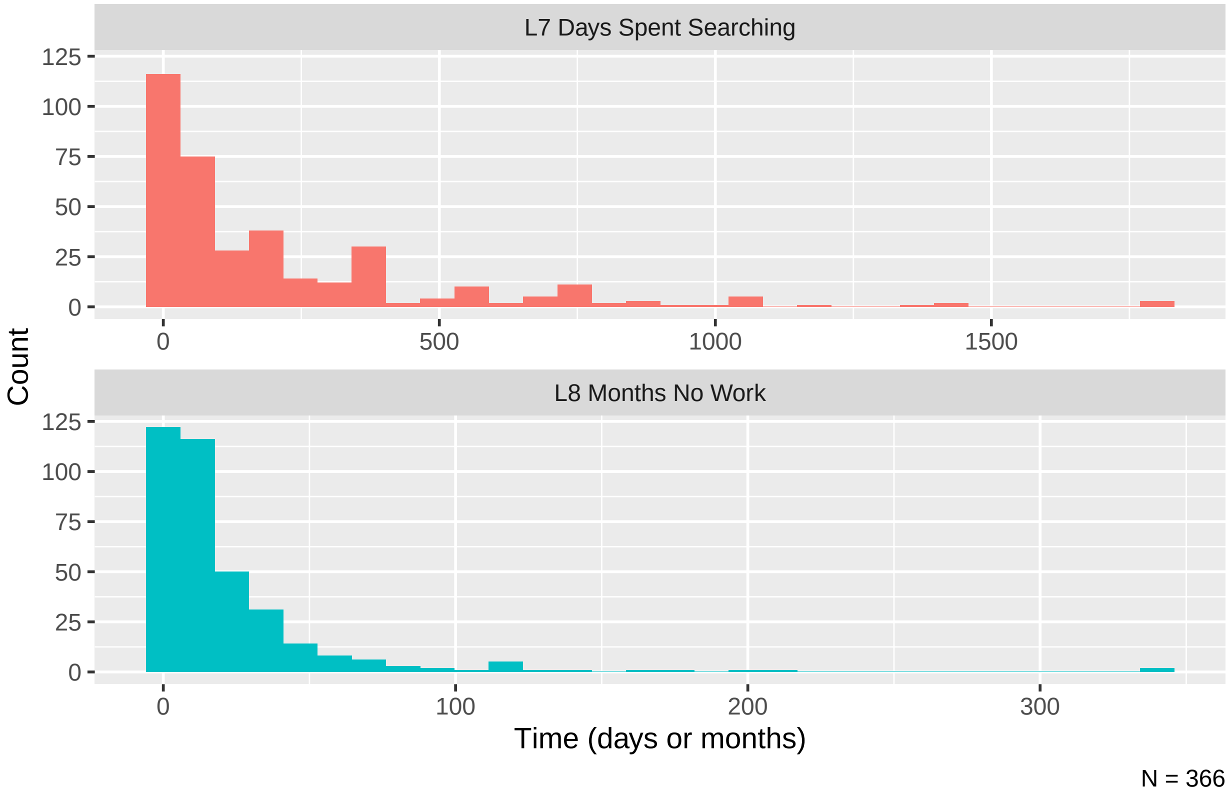
Survey on Consumer Expectations - Job Search Supplement

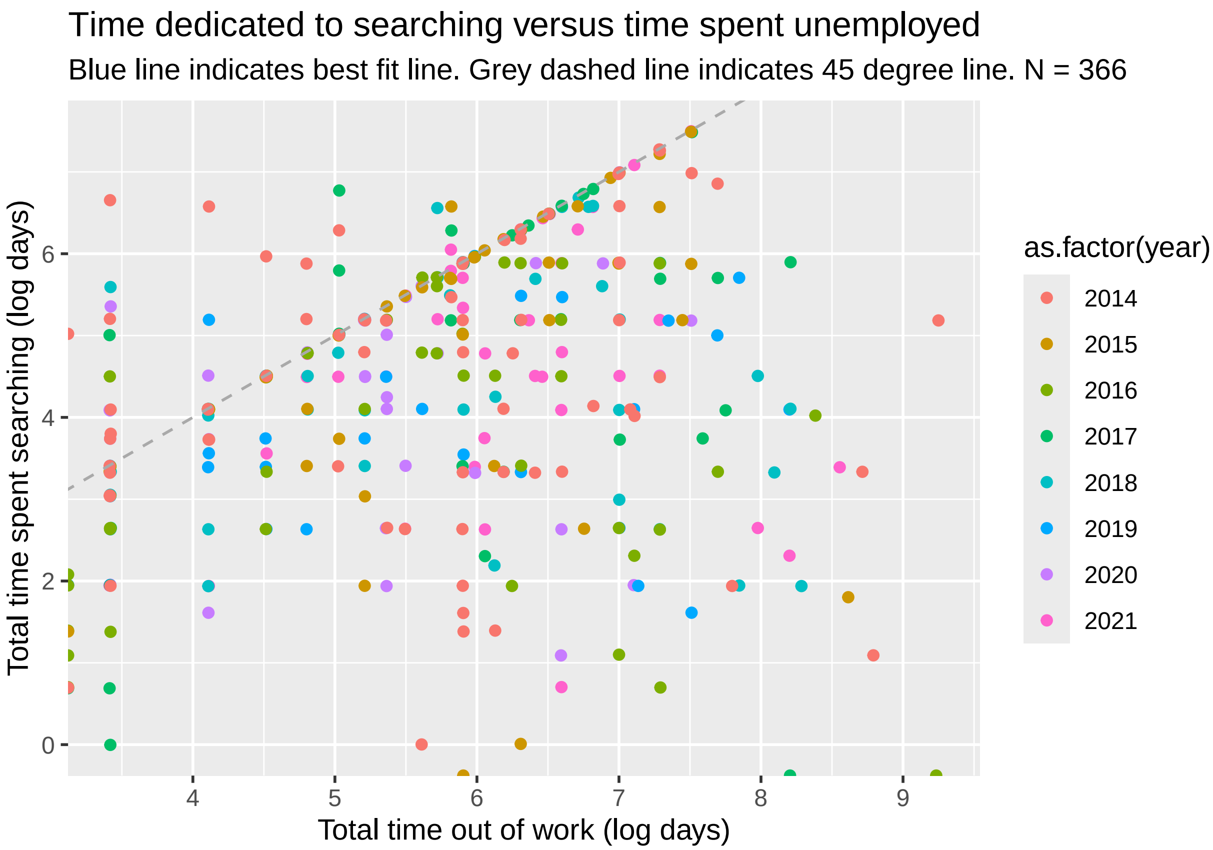
The Federal Reserve Bank of New York compiles the nationally representative Survey on Consumer Expectations annually in October. Since 2013, they have run a Job Search Supplement which includes questions on the time spent searching for work, and unemployment duration. The job search supplement has plenty more questions that we can look at incorporating, listed here. For now, I plot the relationship between time spent searching and time out of work. The table below also indicates the number of people unemployed in the dataset and the number of people unemployed and searching.

Year	N Unemployed	N Unemp & Searching
2014	383	70
2015	321	44
2016	339	46
2017	350	38
2018	354	41
2019	343	32

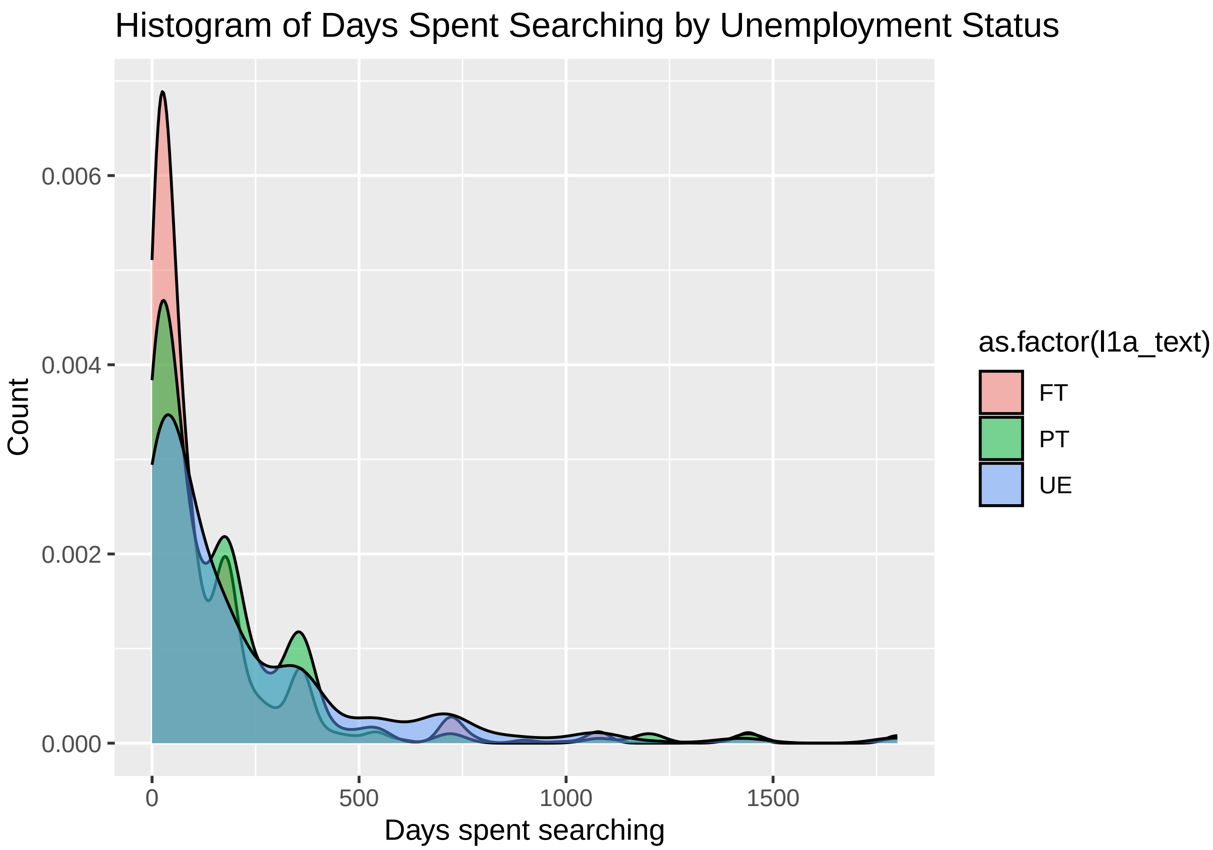
Year	N Unemployed	N Unemp & Searching
2020	304	45
2021	330	50

Histogram of time spent searching and out of work.




Time dedicated to searching versus time spent unemployed
Blue line indicates best fit line. Grey dashed line indicates 45 degree line. N = 366
Total time spent searching (log days)
Total time out of work (log days)
0
2
4
6
4
5
6
7
8
9
as.factor(year)
2014
2015
2016
2017
2018
2019
2020
2021

On the job search



(TBD) American Time Use Survey

The American Time Use Survey gives no indication of time spent in unemployment. It shows how much time is spent searching but does not link to time spent in unemployment. Therefore, I prioritised the datasets above. Krueger & Mueller 2010 impute duration spent unemployed from the ATUS in the following way which could be worth considering.

"Unfortunately, the ATUS interview does not collect information on unemployment duration. Consequently, we derive unemployment duration by taking the unemployment duration reported in the last CPS interview and adding the number of weeks that elapsed between the CPS interview and the ATUS interview. Eighty-six percent of the ATUS interviews were conducted within 3 months of the last CPS interview. For those who were not unemployed at the time of the CPS interview, we impute duration of unemployment by taking half the number of weeks between the CPS and the ATUS interviews. We do not show the weekly LOWESS plot for 13 weeks or less, but simply report the average time allocated to search, as the imputed unemployment duration are quite noisy for those who become unemployed after their last CPS interview."

Table 1: Continuous UE Duration w.o Wage Level Control

	Cont.	Cont. (clipped)	Cont. w. UI	Cont. w. UI (clip
Intercept	1.053***	1.045***	1.053***	1.045***
	(0.006)	(0.004)	(0.006)	(0.004)
Unemployment Duration (Months)	-0.007***	-0.006***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Received Unemployment Compensation			-0.000	0.000
			(0.001)	(0.001)
Exhausted Unemployment Compensation				
dwmosun2				
Female				
Age				
White				
Black				
Mixed				
Married				
High School				
educ_catAssociate's				
Bachelor's Degree				
Postgraduate Degree				
Num.Obs.	4870	4644	4870	4644
R2	0.009	0.012	0.009	0.012
R2 Adj.	0.009	0.012	0.009	0.011
F	46.344		23.169	
RMSE	0.38	0.24	0.38	0.24

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 2: Continuous UE Duration w. Wage Level Control

	Cont.	Cont. (clipped)	Cont. w. UI	Cont. w. UI (clip
Intercept	1.185***	1.131***	1.186***	1.130***
	(0.011)	(0.008)	(0.011)	(0.008)
Hourly Wage of Lost Job	-0.009***	-0.006***	-0.009***	-0.006***
	(0.001)	(0.000)	(0.001)	(0.000)
Unemployment Duration (Months)	-0.007***	-0.006***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Received Unemployment Compensation			-0.000	0.000
			(0.001)	(0.001)
Exhausted Unemployment Compensation				
dwmosun2				
Female				
Age				
White				
Black				
Mixed				
Married				
High School				
educ_catAssociate's				
Bachelor's Degree				
Postgraduate Degree				
Num.Obs.	4870	4644	4870	4644
R2	0.048	0.046	0.048	0.046
R2 Adj.	0.047	0.046	0.047	0.046
F	121.551		81.034	
RMSE	0.37	0.24	0.37	0.24

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 3: Binned UE Duration w.o Wage Level Control

	Disc.	Disc. (clipped)	Disc. w. UI	Disc. w. UI (clipp
Intercept	1.069***	1.055***	1.069***	1.055***
	(0.008)	(0.005)	(0.008)	(0.005)
Unemployment Duration (Binned)	-0.013***	-0.009***	-0.013***	-0.009***
	(0.002)	(0.001)	(0.002)	(0.001)
Received Unemployment Compensation			-0.000	0.000
			(0.001)	(0.001)
Exhausted Unemployment Compensation				
Female				
Age				
White				
Black				
Mixed				
Married				
High School				
educ_catAssociate's				
Bachelor's Degree				
Postgraduate Degree				
Num.Obs.	4870	4644	4870	4644
R2	0.010	0.011	0.010	0.011
R2 Adj.	0.009	0.011	0.009	0.010
F	47.638		23.816	
RMSE	0.37	0.24	0.37	0.24

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 4: Binned UE Duration w. Wage Level Control

	Disc.	Disc. (clipped)	Disc. w. UI	Disc. w. UI (clipped)
Intercept	1.198***	1.139***	1.199***	1.139***
	(0.012)	(0.008)	(0.012)	(0.008)
Hourly Wage of Lost Job	-0.009***	-0.006***	-0.009***	-0.006***
	(0.001)	(0.000)	(0.001)	(0.000)
Unemployment Duration (Binned)	-0.011***	-0.009***	-0.011***	-0.009***
	(0.002)	(0.001)	(0.002)	(0.001)
Received Unemployment Compensation			-0.000	0.000
			(0.001)	(0.001)
Exhausted Unemployment Compensation				
Female				
Age				
White				
Black				
Mixed				
Married				
High School				
educ_catAssociate's				
Bachelor's Degree				
Postgraduate Degree				
Num.Obs.	4870	4644	4870	4644
R2	0.047	0.045	0.047	0.045
R2 Adj.	0.047	0.045	0.047	0.045
F	120.632		80.422	
RMSE	0.37	0.24	0.37	0.24

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

	Heckman Correction	Entropy Balanced Reweight	GLM Reweight
Intercept	1.131***	1.147***	1.143***
	(0.041)	(0.033)	(0.033)
Unemployment Duration (Months)	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)
Female	0.018	0.001	0.001
	(0.014)	(0.011)	(0.011)
Age	-0.007***	-0.002***	-0.002***
	(0.002)	(0.000)	(0.000)
White	-0.162*	-0.027	-0.023
	(0.074)	(0.025)	(0.025)
Black	-0.125*	-0.040	-0.036
	(0.050)	(0.030)	(0.030)
Mixed	-0.054	0.003	0.007
	(0.055)	(0.044)	(0.044)
Married	0.003	0.005	0.004
	(0.011)	(0.011)	(0.011)
High School	-0.014	-0.014	-0.014
	(0.019)	(0.017)	(0.017)
XOeduc_catAssociate's	-0.078		
	(0.064)		
Bachelor's Degree	-0.217	0.054*	0.054*
	(0.165)	(0.023)	(0.023)
Postgraduate Degree	-0.479	0.083+	0.086+
	(0.330)	(0.048)	(0.047)
Inverse Mills Ratio	0.870+		
	(0.479)		
educ_catAssociate's		0.007	0.006
		(0.022)	(0.022)
Num.Obs.	4870	4870	4870
R2	0.893	0.014	0.015
R2 Adj.	0.893	0.012	0.013
F		6.487	6.798
RMSE	0.37	0.37	0.37

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001