# Online Appendix (Not for Publication)

Online Appendix for paper "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias" by Andreas I. Mueller, Johannes Spinnewijn and Giorgio Topa.

## A Survey Questions

### A.1 Survey of Consumer Expectations

#### Question about 12-Month Job Finding Prospect

*What do you think is the percent chance that within the coming 12 months, you will find a job that you will accept, considering the pay and type of work?*

[Ruler & box]

#### Question about 3-Month Job Finding Prospect

*And looking at the more immediate future, what do you think is the percent chance that within the coming 3 months, you will find a job that you will accept, considering the pay and type of work?*

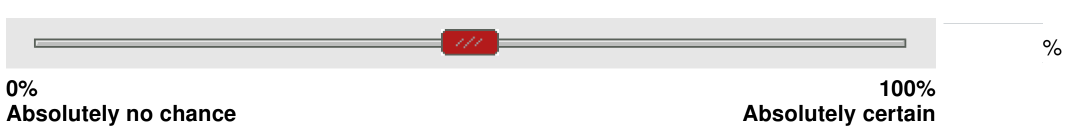
[Ruler & box]

### A.2 Krueger-Mueller Survey

#### Question about 1-Month Job Finding Prospect

*What do you think is the percent chance that you will be employed again within the next 4 weeks?*

*Please move the red button on the bar below to select the percent chance, where 0% means 'absolutely no chance' and 100% means 'absolutely certain'.*




[NB: Initial position on bar is randomized.]

#### Question about Expected Duration

*How many weeks do you estimate it will actually take before you will be employed again?*

________ *Weeks*

# B Descriptive Statistics and Sample Comparisons

This appendix provides additional descriptive statistics for both surveys, for the main samples used in the analysis and a series of sub-samples. Table B1 shows statistics for the SCE, while Table B2 shows statistics for the KM survey. In both tables, column 1 shows the full sample of unemployed, column 2 shows the main sample used in the longitudinal analysis and column 3 shows the main sample used for the analysis of predictive power of beliefs. The samples statistics are very similar across the different main samples, except for the sample in column 3 for the KM survey, which is substantially smaller and somewhat selected. Columns 4 and 5, shows descriptive statistics for the short- and the long-term unemployed. The long-term unemployed in both surveys are older and more female, though little differences exist by education. Finally, columns 6 and 7 show descriptive statistics separately by number of surveys completed. In the SCE, attrition appears to be non-random with respect to education, though survey weights adjust for this. Reassuringly, the monthly job-finding rate is similar across the two samples. In the KM survey, attrition also appears non-random, but again survey weights adjust for this. Moreover, the sample in column 7 with two surveys or more, which is the variation we use for the longitudinal analysis in the KM data, is similar to the full weekly panel, which is representative of the UI population in New Jersey over the survey period.

Table B1: Sample Comparison in the SCE

| (in %) | Unemployed: Main samples (1) | (2) | (3) | Unemployed: Sub-samples of sample (2) (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| High-School Degree or Less | 44.5 | 42.1 | 43.7 | 42.1 | 42.0 | 46.6 | 39.4 |
| Some College Education | 32.4 | 32.6 | 30.3 | 30.3 | 34.5 | 35.9 | 30.6 |
| College Degree or More | 23.1 | 25.4 | 26.0 | 27.6 | 23.5 | 17.5 | 29.9 |
| Ages 20-34 | 25.4 | 23.8 | 21.2 | 29.4 | 19.1 | 29.9 | 20.3 |
| Ages 35-49 | 33.5 | 34.4 | 33.1 | 34.5 | 34.4 | 32.8 | 35.4 |
| Ages 50-65 | 41.1 | 41.7 | 45.7 | 36.1 | 46.5 | 37.4 | 44.2 |
| Female | 59.3 | 58.0 | 55.3 | 53.3 | 62.0 | 60.6 | 56.6 |
| Black | 19.1 | 18.2 | 14.3 | 17.9 | 18.5 | 19.9 | 17.2 |
| Hispanic | 12.5 | 12.2 | 11.0 | 13.2 | 11.4 | 11.3 | 12.8 |
| Monthly job-finding rate | 18.7 | 18.2 | 17.3 | 24.2 | 13.2 | 18.9 | 18.0 |
| # respondents | 948 | 882 | 494 | 513 | 479 | 395 | 487 |
| # respondents w/ > 1 survey | 534 | 477 | 278 | 260 | 239 | 166 | 311 |
| # survey responses | 2,597 | 2,281 | 1,201 | 1,070 | 1,211 | 756 | 1,525 |

*Notes*: All samples are restricted to ages 20-65. Survey weights are used to compute the descriptive statistics. The table shows descriptive statistics in the SCE for (1) all unemployed, (2) all unemployed where the sample was trimmed from observations with inconsistent elicitations between the 3-month and the 12-month probability, and (3) the same sample as in (2) but in addition limited to interviews with at least 3 consecutive monthly follow-up interviews (used to measure 3-month job finding). Sample (2) is the baseline sample in our paper and sample (3) is used for the analysis of the predictive power of beliefs in Section 3.3. Sub-sample (4) consists of unemployed workers of durations 0-6 months and sub-sample (5) of unemployed workers of duration of 7 or more months. Sub-sample (6) consists of those in the main sample (2) with less than 9 interviews over the 12 months of the SCE and sub-sample (7) consists of those in the main sample (2) with at least 9 interviews over the 12 months of the SCE.

Table B2: Sample Comparison in the KM Survey

| (in %) | Weekly panel | Monthly panel: Main samples (1) | (2) | (3) | Monthly panel: Sub-samples of (2) (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|
| High-School Degree or Less | 32.6 | 32.5 | 32.4 | 30.2 | 32.8 | 32.3 | 39.6 | 29.3 |
| Some College Education | 38.0 | 37.4 | 37.4 | 36.7 | 37.0 | 37.5 | 38.0 | 37.1 |
| College Degree or More | 29.5 | 30.1 | 30.2 | 33.0 | 30.2 | 30.2 | 22.4 | 33.7 |
| Ages 20-34 | 35.3 | 38.1 | 38.7 | 21.7 | 44.9 | 35.7 | 55.1 | 31.4 |
| Ages 35-49 | 36.5 | 35.4 | 34.8 | 31.8 | 35.8 | 34.3 | 31.8 | 36.2 |
| Ages 50-65 | 28.2 | 26.5 | 26.5 | 46.4 | 19.3 | 30.0 | 13.1 | 32.4 |
| Female | 48.1 | 48.6 | 47.3 | 42.7 | 46.2 | 47.8 | 50.2 | 46.0 |
| Black | 22.8 | 24.4 | 24.1 | 15.2 | 23.0 | 24.6 | 30.7 | 21.1 |
| Hispanic | 23.6 | 27.5 | 27.8 | 16.9 | 33.7 | 24.9 | 31.8 | 26.0 |
| Monthly job-finding rate | 11.1 | 10.3 | 10.5 | 10.5 | 13.5 | 9.5 | — | 9.2 |
| # respondents | 4,939 | 2,384 | 2,278 | 552 | 397 | 1,884 | 744 | 1,534 |
| # respondents w/ $>1$ survey | 3,835 | 1,422 | 1,296 | 121 | 175 | 1,118 | 0 | 1,296 |
| # survey responses | 27,021 | 4,803 | 4,435 | 734 | 572 | 3,863 | 744 | 3,691 |

*Notes*: All samples are restricted to unemployed workers, ages 20-65. Survey weights are used to compute the descriptive statistics. The table shows descriptive statistics for the full weekly panel of the KM survey, as well as statistics for (1) the full monthly panel with elicitations about job finding , (2) the same sample as in (1) but trimmed for observations with inconsistencies between the elicitation of the 1-month probability and the expected remaining duration, and (3) the same sample as in (2) but in addition limited to interviews with at least 4 consecutive weekly follow-up interviews (used to measure 1-month job finding). Sample (2) is the baseline sample in our paper and sample (3) is used for the analysis of the predictive power of beliefs. Sub-sample (4) consists of unemployed workers of durations 0-6 months and sub-sample (5) of unemployed workers of duration of 7 or more months. Sub-sample (6) consists of those in the main sample (2) with only 1 interview in the monthly panel and sub-sample (2) consists of those with more than 1 interview in the monthly panel.

# C Theoretical Derivations in Conceptual Framework

This appendix provides the derivations underlying the characterizations in Section 3.1.

**Decomposition of Observed Duration Dependence** We first establish the following result linking the expectations at different durations:

$$E_{d+1}(X_{i,\delta}) = E_d\left(X_{i,\delta}\right) - \frac{cov_d\left(X_{i,\delta}, T_{i,d}\right)}{1 - E_d\left(T_{i,d}\right)}. \tag{20}$$

To obtain this expression, we note first that

$$\begin{aligned} E_d(T_{i,d}) &= \int \frac{S_{i,d}}{S_d} T_{i,d} di = \int \frac{S_{i,d}}{S_d} \left( \frac{S_{i,d} - S_{i,d+1}}{S_{i,d}} \right) di, \\ &= \int \frac{S_{i,d} - S_{i,d+1}}{S_d} di = 1 - \frac{S_{d+1}}{S_d}. \end{aligned}$$

where $S_{i,d}$ refers to the survival rate to duration $d$ for individuals of type $i$. We can then re-write

$$\begin{aligned} E_{d+1}(X_{i,\delta}) &= \int \frac{S_{i,d+1}}{S_{d+1}} X_{i,\delta} di = \frac{S_d}{S_{d+1}} \int \frac{S_{i,d}\left(1 - T_{i,d}\right)}{S_d} X_{i,\delta} di \\ &= \frac{1}{1 - E_d\left(T_{i,d}\right)} \left\{ \int \frac{S_{i,d}}{S_d} X_{i,\delta} di - \int \frac{S_{i,d}}{S_d} T_{i,d} X_{i,\delta} di \right\} \\ &= \frac{1}{1 - E_d\left(T_{i,d}\right)} \left\{ \begin{array}{c} \int \frac{S_{i,d}}{S_d} X_{i,\delta} di - \left[ \int \frac{S_{i,d}}{S_d} T_{i,d} di \right] \left[ \int \frac{S_{i,d}}{S_d} X_{i,\delta} di \right] \\ - \left[ \int \frac{S_{i,d}}{S_d} T_{i,d} X_{i,\delta} di - \left[ \int \frac{S_{i,d}}{S_d} T_{i,d} di \right] \left[ \int \frac{S_{i,d}}{S_d} X_{i,\delta} di \right] \right] \end{array} \right\} \\ &= \frac{1}{1 - E_d\left(T_{i,d}\right)} \left\{ E_d\left(X_{i,\delta}\right)\left(1 - E_d\left(T_{i,d}\right)\right) - cov_d\left(T_{i,d}, X_{i,\delta}\right) \right\}, \\ &= E_d\left(X_{i,\delta}\right) - \frac{cov_d\left(T_{i,d}, X_{i,\delta}\right)}{1 - E_d\left(T_{i,d}\right)} \end{aligned}$$

which proves the earlier statement.

The decomposition of the observed duration dependence immediately follows from applying equation (20) to the average job finding. That is,

$$E_d(T_{i,d}) - E_{d+1}\left(T_{i,d+1}\right) = E_d(T_{i,d} - T_{i,d+1}) + \frac{cov_d\left(T_{i,d}, T_{i,d+1}\right)}{1 - E_d\left(T_{i,d}\right)},$$

where

$$\begin{aligned} cov_d\left(T_{i,d}, T_{i,d+1}\right) &= cov_d\left(T_{i,d}, T_{i,d} + \left(T_{i,d+1} - T_{i,d}\right)\right) \\ &= var_d\left(T_{i,d}\right) - cov_d\left(T_{i,d}, T_{i,d} - T_{i,d+1}\right). \end{aligned}$$

**Lower Bound on Heterogeneity** We first note that

$$\begin{aligned}
cov_d\left(Z_{i,d},F_{i,d}\right) &= E_d\left(Z_{i,d}F_{i,d}\right)-E_d\left(Z_{i,d}\right)E_d\left(F_{i,d}\right)\\
&= E_d\left(E_d\left(Z_{i,d}F_{i,d}|T_{i,d}\right)\right)-E_d\left(Z_{i,d}\right)E_d\left(T_{i,d}\right)\\
&= E_d\left(E_d\left(Z_{i,d}\times 1|T_{i,d}\right)\Pr\left(F_{i,d}=1|T_{i,d}\right)+E_d\left(Z_{i,d}\times 0|T_{i,d}\right)\Pr\left(F_{i,d}=0|T_{i,d}\right)\right)-E_d\left(Z_{i,d}\right)E_d\left(T_{i,d}\right)\\
&= E_d\left(E_d\left(Z_{i,d}T_{i,d}|T_{i,d}\right)\right)-E_d\left(Z_{i,d}\right)E_d\left(T_{i,d}\right)\\
&= E_d\left(Z_{i,d}T_{i,d}\right)-E_d\left(Z_{i,d}\right)E_d\left(T_{i,d}\right)\\
&= cov_d\left(Z_{i,d},T_{i,d}\right).
\end{aligned}$$

Now we can use the Cauchy-Schwarz inequality,

$$\begin{aligned}
var_d\left(T_{i,d}\right)var_d\left(Z_{i,d}\right) &\geq cov_d\left(Z_{i,d},T_{i,d}\right)^2\\
&= cov_d\left(Z_{i,d},F_{i,d}\right)^2.
\end{aligned}$$

Hence, we have derived the first lower bound on the variance in job-finding rates,

$$var_d\left(T_{i,d}\right)\geq\frac{cov_d\left(Z_{i,d},F_{i,d}\right)^2}{var_d\left(Z_{i,d}\right)}.$$

We next use Proposition 3 in Morrison and Taubinsky [2019] to derive a second lower bound using two measurements of the beliefs, $Z^1_{i,d}$ and $Z^2_{i,d}$, which requires that both are independently distributed conditional on $T_{i,d}$ (i.e., $Z^1_{i,d}\perp Z^2_{i,d}|T_{i,d}$) and have the same conditional expectation (i.e., $E\left(Z^j_{i,d}|T_{i,d}\right)=\alpha\left(T_{i,d}\right)$). Following the proof in Morrison and Taubinsky [2019], we first note that

$$\begin{aligned}
cov_d\left(Z^1_{i,d},Z^2_{i,d}\right) &= E_d\left(Z^1_{i,d}Z^2_{i,d}\right)-E_d\left(Z^1_{i,d}\right)E_d\left(Z^2_{i,d}\right)\\
&= E_d\left(E_d\left(Z^1_{i,d}Z^2_{i,d}|T_{i,d}\right)\right)-E_d\left(\alpha\left(T_{i,d}\right)\right)E_d\left(\alpha\left(T_{i,d}\right)\right)\\
&= E_d\left(\alpha\left(T_{i,d}\right)^2\right)-E_d\left(\alpha\left(T_{i,d}\right)\right)^2\\
&= var_d\left(\alpha\left(T_{i,d}\right)\right)\text{ and}\\
cov_d\left(F_{i,d},Z^j_{i,d}\right) &= cov_d\left(T_{i,d},\alpha\left(T_{i,d}\right)\right).
\end{aligned}$$

So we can combine these expressions with the Cauchy-Schwarz inequality,

$$var_d\left(T_{i,d}\right)var_d\left(\alpha\left(T_{i,d}\right)\right)\geq cov_d\left(T_{i,d},\alpha\left(T_{i,d}\right)\right)^2$$

to derive the lower bound

$$var_d\left(T_{i,d}\right)\geq\frac{cov_d\left(F_{i,d},Z^1_{i,d}\right)cov_d\left(F_{i,d},Z^2_{i,d}\right)}{cov_d\left(Z^1_{i,d},Z^2_{i,d}\right)}.$$

Note that if the two measurements are positively correlated, conditional on $T_{i,d}$, the lower bound argument continues to hold, since then $cov_d\left(Z^1_{i,d},Z^2_{i,d}\right)\geq var\left(\alpha\left(T_{i,d}\right)\right)$.

**Linear Beliefs Model** For the linear beliefs model, we have

$$E_d\left(Z_{i,d}\right) = b_0 + b_1 E_d\left(T_{i,d}\right) \text{ for any } d.$$

Hence, it trivially follows that

$$b_1 = \frac{E_{d+1}\left(Z_{i,d+1}\right) - E_d\left(Z_{i,d}\right)}{E_{d+1}\left(T_{i,d+1}\right) - E_d\left(T_{i,d}\right)}.$$

The variance in job-finding probabilities is thus identified by

$$var_d\left(T_{i,d}\right) = \frac{1}{b_1} \times cov_d\left(F_{i,d}, Z_{i,d}\right) = \frac{E_{d+1}\left(T_{i,d+1}\right) - E_d\left(T_{i,d}\right)}{E_{d+1}\left(Z_{i,d+1}\right) - E_d\left(Z_{i,d}\right)} \times cov_d\left(F_{i,d}, Z_{i,d}\right)$$

We also note that with two elicitations $Z_{i,d}^1$ and $Z_{i,d}^2$ that are both some linear transformation of $T_{i,d}$,

$$Z_{i,d}^j = b_0^j + b_1^j T_{i,d} + \varepsilon_{i,d}^j,$$

our earlier lower bound becomes tight,

$$var_d\left(T_{i,d}\right) = \frac{cov_d\left(F_{i,d}, Z_{i,d}^1\right) cov_d\left(F_{i,d}, Z_{i,d}^2\right)}{cov_d\left(Z_{i,d}^1, Z_{i,d}^2\right)}.$$

But this is only the case when the error terms are independently distributed. The covariance ratio becomes a lower bound when the error terms are positively correlated.

**Persistence in Job Finding** For the linear beliefs model, it naturally follows that

$$cov_{d+1}\left(F_{i,d+1}, Z_{i,d}\right) = b_1 cov_{d+1}\left(T_{i,d+1}, T_{i,d}\right).$$

This term relates to the dynamic selection term contributing to the observed duration dependence in job finding, but is backward-looking rather than forward-looking. We characterize this relationship in case the job finding probability consists of a permanent term and a random transitory term, independently drawn every period,

$$T_{i,d} = T_i + \tau_{i,d}.$$

Here, the forward-looking covariance simplifies to the variance in the persistent component,

$$cov_d\left(T_{i,d+1}, T_{i,d}\right) = var_d\left(T_i\right).$$

The backward-looking covariance can be re-expressed as

$$\begin{aligned}
cov_{d+1}\left(T_{i,d+1},T_{i,d}\right) &= cov_{d+1}\left(T_i+\tau_{i,d+1},T_i+\tau_{i,d}\right) \\
&= E_{d+1}\left(\left(T_i+\tau_{i,d+1}\right)\left(T_i+\tau_{i,d}\right)\right)-E_{d+1}\left(T_i+\tau_{i,d+1}\right)E_{d+1}\left(T_i+\tau_{i,d}\right) \\
&= E_{d+1}\left[E_{d+1}\left(\left(T_i+\tau_{i,d+1}\right)\left(T_i+\tau_{i,d}\right)|T_i\right)\right] \\
&\quad -E_{d+1}\left[E_{d+1}\left(T_i+\tau_{i,d+1}|T_i\right)\right]E_{d+1}\left[E_{d+1}\left(T_i+\tau_{i,d}|T_i\right)\right] \\
&= E_{d+1}\left(T_i^2+T_i\tau_{i,d}\right)-E_{d+1}\left(T_i\right)E_{d+1}\left(T_i+\tau_{i,d}\right) \\
&= var_{d+1}(T_i)+cov_{d+1}(T_i,\tau_{i,d})
\end{aligned}$$

Here, we can apply the characterization of the conditional expectation in (20) to re-express

$$\begin{aligned}
cov_{d+1}(T_i,\tau_{i,d}) &= E_{d+1}\left(T_i\tau_{i,d}\right)-E_{d+1}\left(T_i\right)E_{d+1}\left(\tau_{i,d}\right) \\
&= E_d\left(T_i\tau_{i,d}\right)-\frac{cov_d\left(T_i\tau_{i,d},T_{i,d}\right)}{1-E_d\left(T_{i,d}\right)}-E_{d+1}\left(T_i\right)\left[E_d\left(\tau_{i,d}\right)-\frac{cov_d\left(\tau_{i,d},T_{i,d}\right)}{1-E_d\left(T_{i,d}\right)}\right] \\
&= -\frac{E_d\left(T_i\tau_{i,d}^2\right)}{1-E_d\left(T_{i,d}\right)}+E_{d+1}\left(T_i\right)\frac{var_d\left(\tau_{i,d}\right)}{1-E_d\left(T_{i,d}\right)} \\
&= -\frac{1}{1-E_d\left(T_{i,d}\right)}\left[E_d\left(T_i\tau_{i,d}^2\right)-E_d(T_i)E\left(\tau_{i,d}^2\right)-\left(E_{d+1}\left(T_i\right)-E_d\left(T_i\right)\right)var_d\left(\tau_{i,d}\right)\right] \\
&= -\frac{1}{1-E_d\left(T_{i,d}\right)}\left[cov\left(T_i,\tau_{i,d}^2\right)-\left(E_{d+1}\left(T_i\right)-E_d\left(T_i\right)\right)var_d\left(\tau_{i,d}\right)\right] \\
&= \frac{E_{d+1}\left(T_i\right)-E_d\left(T_i\right)}{1-E_d\left(T_{i,d}\right)}var_d\left(\tau_{i,d}\right) \\
&= -\frac{var_d\left(T_i\right)var_d\left(\tau_{i,d}\right)}{\left[1-E_d\left(T_{i,d}\right)\right]^2}
\end{aligned}$$

where we have used

$$\begin{aligned}
cov_d\left(T_i\tau_{i,d},T_i+\tau_{i,d}\right) &= E_d\left(T_i^2\tau_{i,d}\right)+E_d\left(T_i\tau_{i,d}^2\right)-E_d\left(T_i\tau_{i,d}\right)E_d\left(T_i\right)-E_d\left(T_i\tau_{i,d}\right)E_d\left(\tau_{i,d}\right) \\
&= E_d\left(T_i\tau_{i,d}^2\right), \\
E_{d+1}\left(T_i\right)-E_d\left(T_i\right) &= -\frac{cov_d\left(T_i,T_i\right)}{1-E_d\left(T_{i,d}\right)} \text{ and} \\
cov\left(T_i,\tau_{i,d}^2\right) &= E_d\left(T_i\tau_{i,d}^2\right)-E_d(T_i)E\left(\tau_{i,d}^2\right)=0.
\end{aligned}$$

Hence, we have

$$cov_{d+1}\left(T_{i,d+1},T_{i,d}\right)=var_{d+1}(T_i)-\frac{var_d\left(T_i\right)var_d\left(\tau_{i,d}\right)}{\left[1-E_d\left(T_{i,d}\right)\right]^2}$$

while instead

$$cov_{d+1}\left(T_{i,d+1},T_{i,d+1}\right)=var_{d+1}(T_i)+var_{d+1}\left(\tau_{i,d+1}\right)=var_{d+1}(T_i)+var_d\left(\tau_{i,d}\right)$$

Putting the two expressions together, we obtain

$$cov_{d+1}\left(T_{i,d+1}, T_{i,d+1}\right) - cov_{d+1}\left(T_{i,d+1}, T_{i,d}\right) = var_d\left(\tau_{i,d}\right)\left[1 - \frac{var_d\left(T_i\right)}{\left[1 - E_d\left(T_{i,d}\right)\right]^2}\right],$$

which shows that the difference in the contemporaneous covariance and lagged covariance is increasing in the variance in transitory shocks, but decreasing in the variance in the permanent shocks. Combined with

$$cov_d\left(T_{i,d}, T_{i,d}\right) = var_d\left(T_i\right) + var_d\left(\tau_{i,d}\right)$$

and, for the linear beliefs model,

$$\begin{aligned} cov_{d+1}\left(F_{i,d+1}, Z_{i,d}\right) &= b_1 cov_{d+1}\left(T_{i,d+1}, T_{i,d}\right) \\ cov_{d+1}\left(F_{i,d+1}, Z_{i,d+1}\right) &= b_1 cov_{d+1}\left(T_{i,d+1}, T_{i,d+1}\right) \\ cov_{d+1}\left(F_{i,d}, Z_{i,d}\right) &= b_1 cov_d\left(T_{i,d}, T_{i,d}\right) \end{aligned}$$

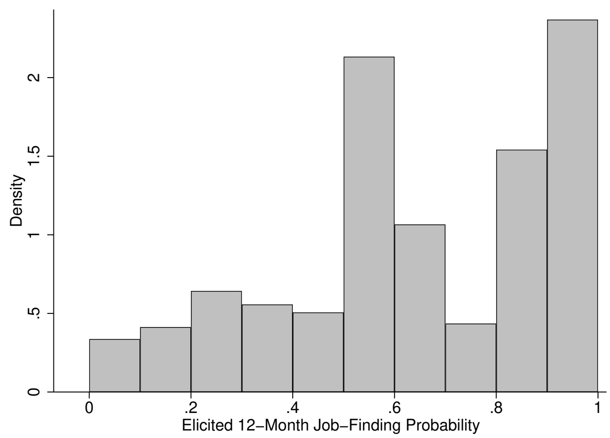
we can use the difference in contemporaneous and lagged covariance to get an estimate of the variance in transitory shocks.

# D Additional Empirical Results

In this appendix, we provide additional results on the empirical analysis that we refer to in the paper for both the SCE and the KM survey. More precisely, we provide (1) additional details on the distributions of elicitations and their correlations, (2) additional details and a series of robustness checks on the predictive power of beliefs and the non-parametric lower bounds, (3) additional details on the biases in beliefs and a detailed robustness analysis of the longitudinal changes in beliefs, (4) evidence on the relationship of beliefs, job search effort and reservation wages in the KM data, and (5) evidence on the relationship between beliefs and indicators of labor market tightness in the SCE.

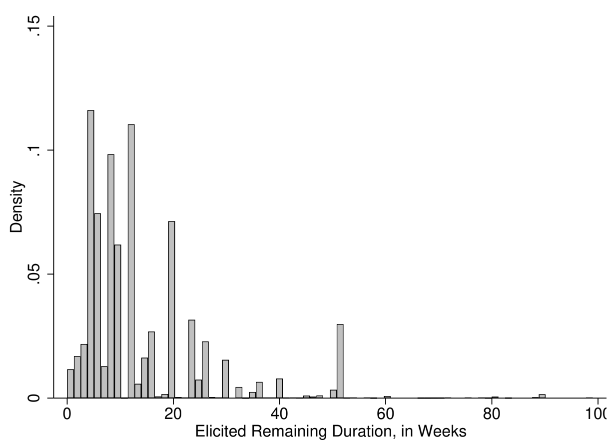
## D.1 Elicited Beliefs about Job Finding

Figure D1: Histogram of the Elicited 12-Month Job-Finding Probability in the SCE



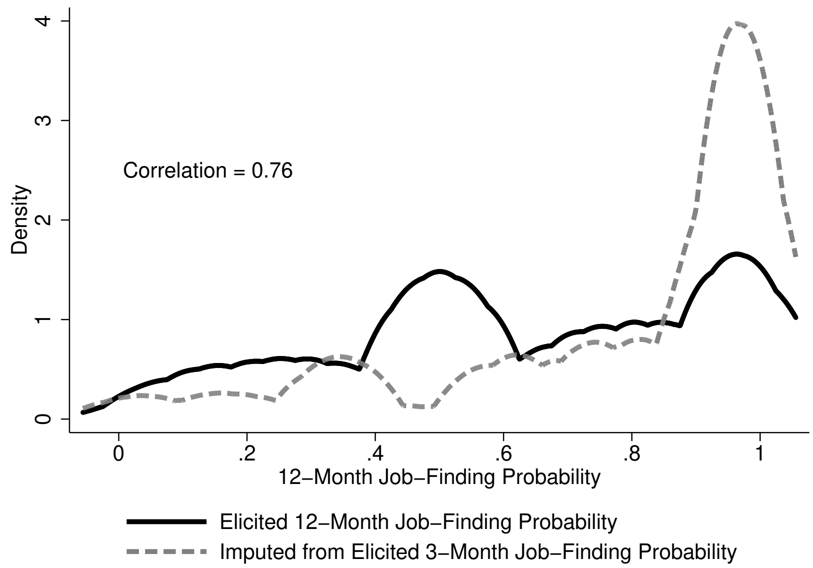

*Notes:* Survey weights are used and the sample is restricted to unemployed workers, ages 20-65.

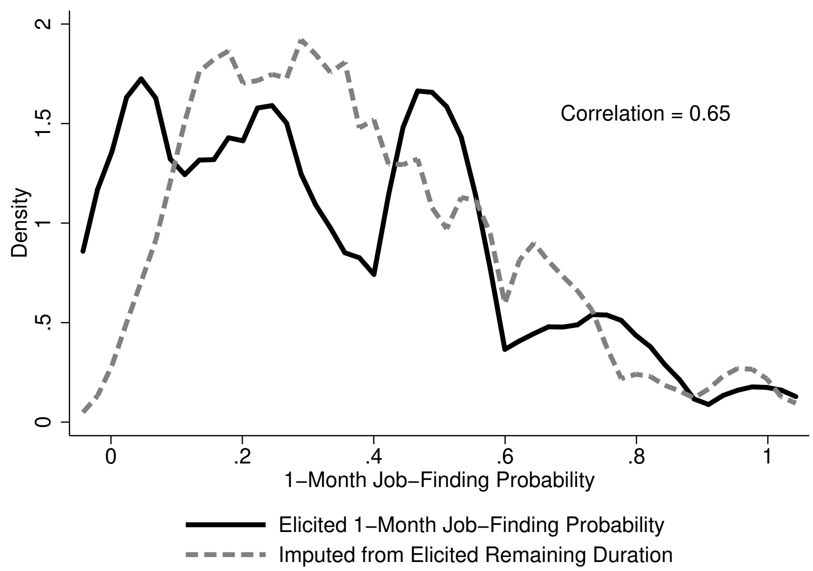
Figure D2: Histogram of Elicitations of the Expected Remaining Duration in the KM survey




*Notes:* Survey weights are used and the sample is restricted to unemployed workers, ages 20-65.

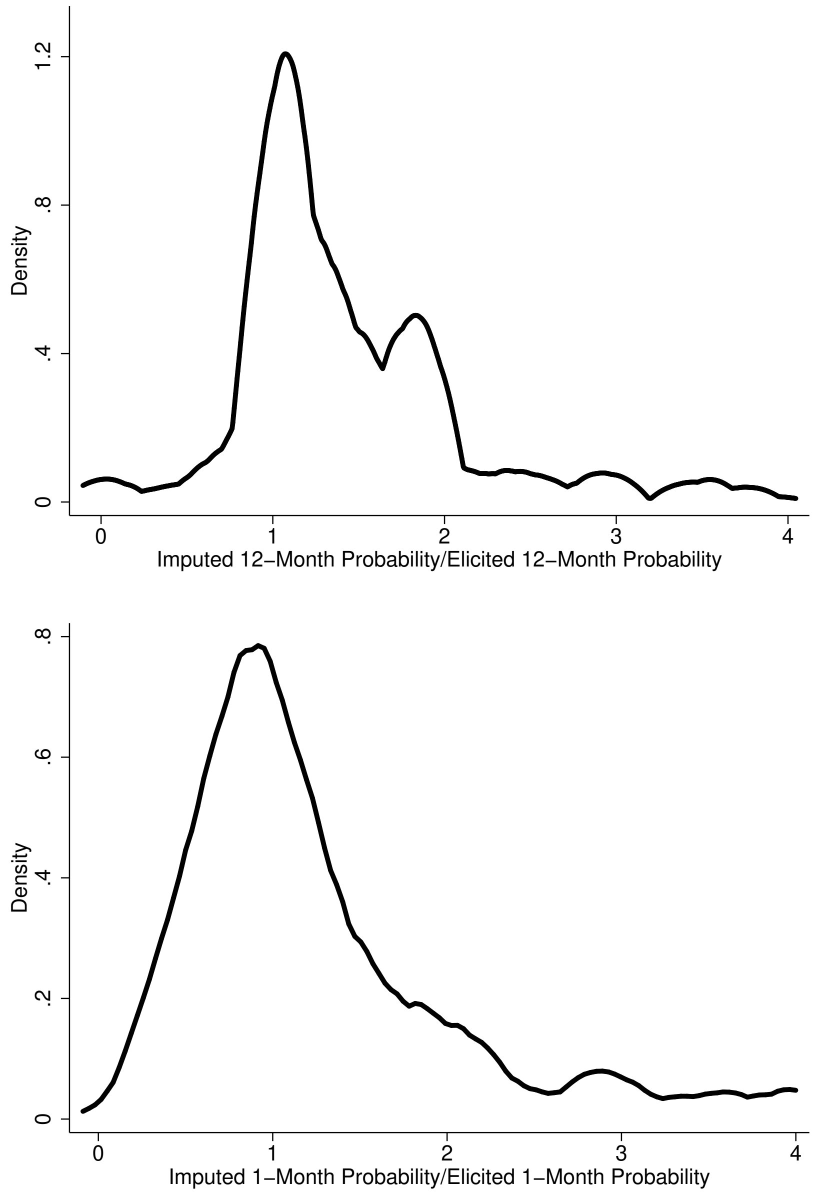
Figure D3: Comparison of Kernel Density Estimates for Alternative Forms of Elicitations




*Notes:* Survey weights are used and the samples are restricted to unemployed workers, ages 20-65. The 12-month probability imputed from the elicited 3-month job-finding probability is computed as $1-(1-Z^3)^4$, where $Z^3$ is the elicited 3-month job-finding probability. The 1-month probability imputed from the elicited remaining duration is computed as $1-(1-\frac{1}{wks})^4$, where $wks$ is the elicited remaining duration unemployed.

Figure D4: Ratio of Imputed Probabilities and Elicited Probabilities based on Alternative Forms of Elicitations in the SCE (top panel) and KM Survey (bottom panel)




*Notes:* Survey weights are used and the samples are restricted to unemployed workers, ages 20-65. The 12-month probability imputed from the elicited 3-month job-finding probability is computed as $1-(1-Z^3)^4$, where $Z^3$ is the elicited 3-month job-finding probability. The 1-month probability imputed from the elicited remaining duration is computed as $1-(1-\frac{1}{wks})^4$, where $wks$ is the elicited remaining duration unemployed.

## D.2 Job Finding Beliefs and Outcomes

In this section, we provide additional details on the analysis of the predictive value of beliefs, a series of robustness checks of the findings in Table 2 in the paper as well as the results for the KM survey. The section also shows additional estimates of the non-parametric lower bounds in the SCE as well as the lower bound estimates for the KM survey.

Table D1: Regressions of Realized 3-Month Job Finding on Elicitations (Showing Controls)

| Dependent Variable: Indicator Variable for Job Finding | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Elicited 3-Month Job-Finding Probability | 0.586*** | | 0.464*** | 0.501*** |
| | (0.073) | | (0.069) | (0.092) |
| Elicited 3-Month Job-Finding Probability x Long-Term Unemployed | | | | -0.258* |
| | | | | (0.142) |
| Long-Term Unemployed | | | | -0.078 |
| | | | | (0.094) |
| Female | | -0.134** | -0.114** | -0.071 |
| | | (0.055) | (0.050) | (0.046) |
| Age | | 0.017 | 0.018 | 0.019 |
| | | (0.017) | (0.015) | (0.013) |
| Age*Age | | -0.0003 | -0.0003* | -0.0003* |
| | | (0.0002) | (0.0002) | (0.0001) |
| High-School Degree | | 0.216* | 0.142 | 0.111 |
| | | (0.130) | (0.124) | (0.101) |
| Some College | | 0.136 | 0.098 | 0.075 |
| | | (0.121) | (0.120) | (0.095) |
| College Degree | | 0.117 | 0.054 | 0.031 |
| | | (0.121) | (0.119) | (0.096) |
| Post-Graduate Education | | 0.139 | 0.059 | 0.034 |
| | | (0.128) | (0.126) | (0.105) |
| Other Education | | 0.487* | 0.388 | 0.329* |
| | | (0.267) | (0.261) | (0.197) |
| HH income: \$30,000-\$59,999 | | 0.127** | 0.111** | 0.095** |
| | | (0.055) | (0.053) | (0.048) |
| HH income: \$60,000-\$100,000 | | 0.182*** | 0.189*** | 0.150** |
| | | (0.058) | (0.057) | (0.060) |
| HH income: \$100,000+ | | 0.151* | 0.156** | 0.113 |
| | | (0.079) | (0.072) | (0.082) |
| Race/Ethnicity: Hispanic | | -0.022 | -0.049 | -0.062 |
| | | (0.073) | (0.067) | (0.063) |
| Race/Ethnicity: African-American | | 0.157** | 0.087 | 0.088 |
| | | (0.078) | (0.078) | (0.070) |
| Race/Ethnicity: Asian | | 0.081 | 0.137 | 0.157* |
| | | (0.100) | (0.084) | (0.092) |
| Race/Ethnicity: Other | | -0.145* | -0.131* | -0.125** |
| | | (0.076) | (0.073) | (0.060) |
| Constant | 0.121*** | 0.171 | -0.104 | -0.062 |
| | (0.040) | (0.402) | (0.363) | (0.329) |
| Observations | 1,201 | 1,201 | 1,201 | 1,201 |
| $R^2$ | 0.131 | 0.148 | 0.218 | 0.259 |

*Notes*: The table shows the results in Table 2 in the paper with controls. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate stat. significance at the *0.1, **0.05 and ***0.01 level.

Table D2: Regressions of Realized 3-Month Job Finding on Elicited 3-Month Job-Finding Probabilities: Robustness Checks

**Panel A.** Linear Regressions w/o Demographic Controls

| Dep. Var.: Indicator Variable for Job Finding | (0) | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Elicited Job-Finding | 0.586*** | 0.710*** | 0.587*** | 0.553*** | 0.624*** | 0.444*** | 0.653*** | 0.547*** | 0.567*** | 0.503*** | 0.345*** |
| Probability | (0.073) | (0.088) | (0.073) | (0.075) | (0.079) | (0.097) | (0.143) | (0.074) | (0.106) | (0.172) | (0.109) |
| Cohort & Time F.E. | | | | | | | | x | | | |
| Observations | 1,201 | 1,200 | 989 | 1,356 | 1,066 | 408 | 295 | 1,201 | 369 | 184 | 648 |
| $R^2$ | 0.131 | 0.125 | 0.159 | 0.121 | 0.151 | 0.076 | 0.184 | 0.208 | 0.137 | 0.094 | 0.046 |

**Panel B.** Linear Regressions w/ Demographic Controls

| Dep. Var.: Indicator Variable for Job Finding | (0) | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Elicited Job-Finding | 0.464*** | 0.580*** | 0.458*** | 0.445*** | 0.493*** | 0.322*** | 0.428*** | 0.429*** | 0.450*** | 0.499*** | 0.235** |
| Probability | (0.069) | (0.086) | (0.070) | (0.071) | (0.073) | (0.101) | (0.121) | (0.072) | (0.108) | (0.168) | (0.101) |
| Demogr. Controls | x | x | x | x | x | x | x | x | x | x | x |
| Cohort & Time F.E. | | | | | | | | x | | | |
| Observations | 1,201 | 1,200 | 989 | 1,356 | 1,066 | 408 | 295 | 1,201 | 369 | 184 | 648 |
| $R^2$ | 0.218 | 0.213 | 0.253 | 0.205 | 0.236 | 0.198 | 0.322 | 0.291 | 0.257 | 0.333 | 0.203 |

*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers, ages 20-65. Column (0) show the baseline estimates; in column (1), we instrument the 3-month probability with the 12-month probability (or rather its equivalent in terms of a 3-month probability); in column (2), we exclude answers equal to 50%; in column (3), we do not trim the sample for inconsistent answers between the 3- and 12-month elicitation (i.e., where the 3-month probability was higher than the 12-month probability); in column (4), we only include one unemployment spell for each person; in column (5), we include only one observation for each person (=the first observation in the survey); in column (6), we only include those individuals who entered the survey as employed and became unemployed during the survey period; in column (7) we control for survey-cohort, calendar-year and month-of-the year fixed effects; in column (8), we only include those unemployed for 3 months or less; in column (9), we only include those unemployed for 4-6 months; and in column (10), we only include those unemployed for 7 months or more. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

Table D3: Linear Regressions of Realized Job Finding on Elicitations (KM Survey)

**Panel A.** Elicited 1-Month Job-Finding Probability

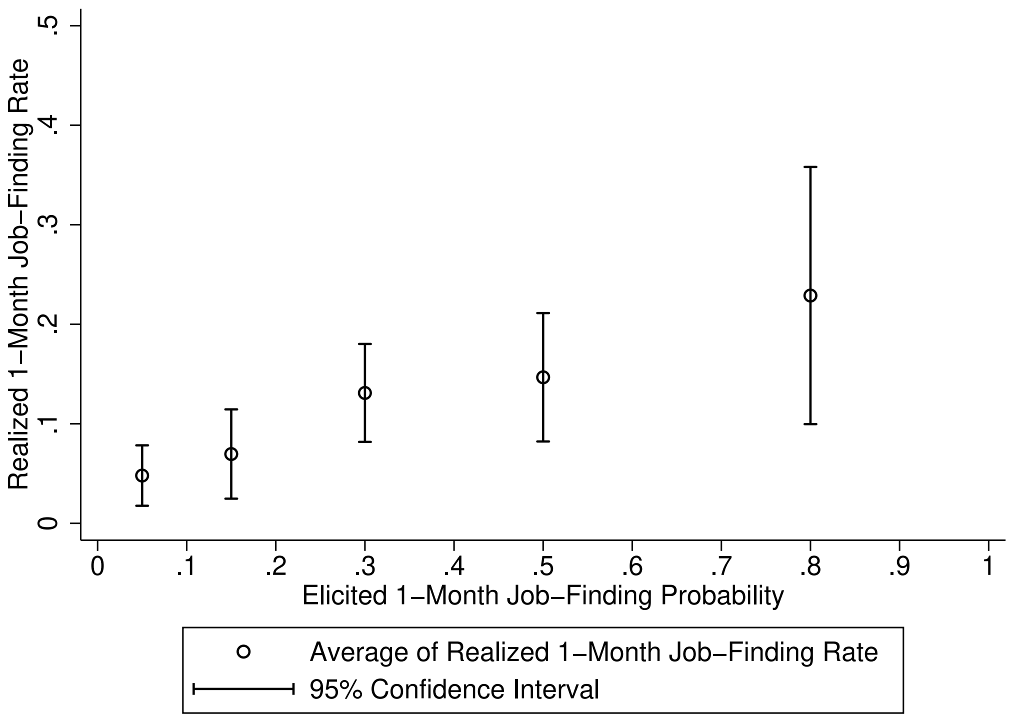
| Dependent Variable: Indicator Variable for Realized 1-Month Job Finding | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Elicited 1-Month Job-Finding Probability | 0.260** | | 0.266*** | 0.382*** |
| | (0.109) | | (0.094) | (0.131) |
| Elicited 1-Month Job-Finding Probability x Long-Term Unemployed | | | | -0.283 |
| | | | | (0.180) |
| Long-Term Unemployed | | | | 0.077* |
| | | | | (0.045) |
| Controls | | | x | x |
| Observations | 650 | 650 | 650 | 650 |
| $R^2$ | 0.039 | 0.189 | 0.224 | 0.234 |

**Panel B.** Elicited Expected Remaining Duration (Inverted)

| Dependent Variable: Indicator Variable for Realized 1-Month Job Finding | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Elicited Remaining Duration (Inverted) | 0.402 | | 0.294 | 0.513 |
| | (0.178)** | | (0.090)*** | (0.140)*** |
| Elicited Remaining Duration (Inverted) x Long-Term Unemployed | | | | -0.493 |
| | | | | (0.157)*** |
| Long-Term Unemployed | | | | 0.145 |
| | | | | (0.056)*** |
| Controls | | | x | x |
| Observations | 650 | 650 | 650 | 650 |
| $R^2$ | 0.080 | 0.189 | 0.223 | 0.249 |

*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers in the KM survey, ages 20-65, with 4 consecutive weekly interviews following the belief question. For the purpose of comparability across the columns in the table, the samples are restricted to the same number of observations, i.e., where all control variables and both elicitations are observed. The elicited expected remaining duration of unemployment (in weeks) is inverted to make it comparable to a 1-month job-finding probability, computed as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration. Controls are dummies for gender, race, ethnicity, household income brackets (4), educational attainment (6), and age and age squared. Long-term unemployment is defined as a duration of unemployment of 60 weeks or more at the beginning of the survey. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

Figure D5: Averages of Realized Job-Finding Rates, by Bins of Elicited Probabilities (KM Survey)




*Notes:* Survey weights are used and the sample is restricted to unemployed workers in the KM survey, ages 20-65, with 4 consecutive weekly interviews following the belief question. The figures show averages of the realized 1-month job-finding rate for five bins of elicited 1-month job-finding probabilities (0-0.1, 0.1-0.2, 0.2-0.4, 0.4-0.6, 0.6-1). For the purpose of comparability across the columns in the table, the samples are restricted to the same number of observations, i.e., where all control variables and both elicitations are observed. The elicited expected remaining duration of unemployment (in weeks) is inverted to make it comparable to a 1-month job-finding probability, computed as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration.

Table D4: Lower Bounds: Additional Estimates (SCE)

| Lower-bound measure based on | Full sample | | Residualized (full sample) | | Short-term unemployed | |
|---|---|---|---|---|---|---|
| | Value | S.e. | Value | S.e. | Value | S.e. |
| ... 3-month elicitations | 0.032 | (0.009) | 0.016 | (0.005) | 0.031 | (0.012) |
| ... 12-month elicitations | 0.028 | (0.008) | 0.013 | (0.004) | 0.016 | (0.009) |
| ... both elicitations | 0.038 | (0.010) | 0.019 | (0.006) | 0.029 | (0.013) |
| ... controls | 0.036 | (0.009) | 0.020 | (0.006) | 0.043 | (0.012) |
| ... controls and 3-month elicitations | 0.053 | (0.010) | 0.030 | (0.007) | 0.059 | (0.014) |
| ... controls and both elicitations | 0.054 | (0.010) | 0.031 | (0.007) | 0.060 | (0.014) |

*Notes*: Standard errors are bootstrapped with 2,000 replications. The lower bounds based on 3- and/or 12-month elicitations are computed according to equations (3) and (4), respectively. The lower bounds based on controls (and elicitations) are the variance of the predicted value of a linear regression of a dummy for realized job finding on controls (and the elicited 3- and/or 12-month job-finding probability). For the results in the second column, in a first stage, beliefs (in rows 1 to 3) and the predicted values (in rows 4 to 6) are regressed on three dummies for unemployment duration (4-6 months, 7-12 months and 13+ months) and then, in a second stage, the residual of the regression is used to compute the lower bounds. The sample in the third column includes only those unemployed with a duration of 3 months or less.

Table D5: Lower Bounds Based on KM Survey

| Lower bound measure based on: | Value | S.e. |
|---|---|---|
| ... 1-month elicitations | 0.0039 | (0.0035) |
| ... elicited remaining duration (inverted) | 0.0081 | (0.0086) |
| ... both elicitations | 0.0085 | (0.0076) |
| ... controls | 0.0192 | (0.0137) |
| ... controls and both elicitations | 0.0233 | (0.0143) |

*Notes*: The lower bounds based on elicitations are computed according to equations (3) and (4). The lower bounds based on controls (and elicitations) are the variance of the predicted value of a linear regression of a dummy for realized job finding on controls (and the elicited 1-month probabilities). The elicited expected remaining duration of unemployment (in weeks) is inverted to make it comparable to a 1-month job-finding probability, computed as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration. Standard errors are bootstrapped with 2,000 replications.

## D.3 Biases in Beliefs

### D.3.1 Average Bias by Duration

The following table shows the statistics in the SCE underlying the Figure 3 in the paper. It also shows the corresponding statistics in the KM survey.

Table D6: Comparison of Perceived and Realized Job Finding

| | Elicited Job-Finding Probability | Realized Job-Finding Rate | Sample Size |
|---|---|---|---|
| **Panel A.** SCE (3-month horizon) | | | |
| Full Sample | 0.491 (0.014) | 0.408 (0.022) | 1,201 |
| Duration 0-3 months | 0.616 (0.028) | 0.642 (0.037) | 369 |
| Duration 4-6 months | 0.529 (0.031) | 0.472 (0.050) | 184 |
| Duration 7-12 months | 0.530 (0.024) | 0.333 (0.045) | 198 |
| Duration 13+ months | 0.354 (0.015) | 0.221 (0.027) | 450 |
| **Panel B.** KM Survey (1-month horizon) | | | |
| Full sample | 0.256 (0.019) | 0.105 (0.022) | 734 |
| Duration 0-6 months | 0.256 (0.042) | 0.135 (0.043) | 79 |
| Duration 7-12 months | 0.283 (0.031) | 0.116 (0.048) | 158 |
| Duration 13+ months | 0.232 (0.028) | 0.076 (0.022) | 497 |

*Notes*: Survey weights are used for all statistics. All samples are restricted to unemployed workers, ages 20-65. The KM sample is further restricted to interviews where the belief questions were administered. Standard errors are in parentheses. Duration refers to self-reported duration in the SCE and duration of weeks of benefit receipt in the KM survey. The SCE sample for this table is restricted to individuals with 4 consecutive interviews. Actual job finding is measured in the SCE as the fraction of individuals who reported being employed in month t+1, t+2 or t+3, where t is the month of the interview where the belief was reported. The KM sample is restricted to those who have not accepted a job in the same or any previous interviews and are not working at the time of the interview. Actual job finding in the KM survey is measured as the fraction accepting a job offer or working in an interview at any point in the 31 days following the interview where the belief was reported.

### D.3.2 Job Finding Beliefs by Duration: Robustness

In this section, we provide additional details on the robustness analysis of our longitudinal results in Table 4 and Figure 4 in the main paper.

We probe the robustness of the finding that beliefs are not revised downward in several ways and also evaluate potential forces that may underlie the (weakly) increasing beliefs about job-finding probabilities. First, we check whether the results in column 4 of Table 4 hold for other measures of perceived job finding. In the KM survey, we find that the expected remaining duration decreases with duration of unemployment when controlling for individual fixed effect. This is obviously consistent with an increasing probability over the spell of unemployment as reported in Table 4. For the purpose of comparison with the probability question, we take the inverse of the expected duration question and convert it into a 4-week probability, assuming that the probability is constant over the spell of unemployment (see footnote 11 for details). Table D8 reports these results. We find that the coefficient is 0.013, which is close to the estimate based on the probability question (0.022). Using the 12-month probabilities in the SCE, the coefficient on unemployment duration is negative but insignificant and very close to zero with an estimate of $-0.0027$ (0.0065). The point estimate implies that the 12-month probability decreases by 3.2 percentage points over a 12-month period, which is almost trivial.

The first columns in Tables D7 and D8 report results where we exclude answers of 50 percent, results where we exclude answers of 100 percent, results where we do not trim outlier answers as discussed further above, and results where we use self-reported duration of unemployment as the independent variable. While individuals increase their perceived job-finding probability as they approach re-employment, the result remains if we exclude individuals who find and accept a job within the next 4 weeks in the KM survey. Neither is the estimate affected when we exclude individuals who reported a job offer in a previous interview but did not accept it (see Table D8). Across all these different specifications, the results are very similar.

We also find that our results are robust to controlling for changes in aggregate labor market conditions during our sample period. For the SCE, which uses a rotating panel, controlling for changes in the national, state unemployment rates or quarterly GDP growth has little effect on our estimate of the duration dependence in perceived job-finding probabilities. More importantly, we also run a specification with individual fixed effects and fixed effects for calendar time, and find that the results are not affected. Even though calendar time and duration are collinear in a given spell, this model is identified in the presence of multiple unemployment spells per person. The results are very similar to our baseline specification, demonstrating that aggregate time trends are unlikely to drive our longitudinal results on perceived job finding. Note that, for the KM survey, the sample period coincides for all job seekers, so calendar time and time spent unemployed are collinear and thus it is problematic to include the state or local unemployment rate into the fixed effect regression. As discussed in Krueger and Mueller [2011], however, the unemployment rate in NJ was nearly constant over the period of the survey (October 2009 through April 2010) between 9.5 and 9.8 and did not drop below 9.4 until August 2011, so it seems unlikely that people perceived the job market to improve over the sample period.

Finally, one may be concerned that the beliefs are not elicited for individuals who get discouraged and drop out of the labor and that as a consequence our longitudinal estimates are upward biased.

While this is a potential concern in the SCE where individuals were asked the belief questions only when unemployed, we'd like to start by emphasizing that this is not the case for the KM survey, where individuals were followed and their perceptions were elicited *independently* of whether they indicated searching for a job or not. To further assess this issue in the SCE, we split the sample into those unemployment spells that were interrupted by or ended with a transition into out of the labor force. We find a negligent longitudinal decline for those spells with a UN transition of 0.0022 percentage points per month (or 2.6 percentage points per year). We further assess this issue by imputing the longitudinal decline for the months where the individual was reported to be out of the labor force based on the longitudinal decline in the months unemployed. We then re-estimate our regression with spell fixed effects but now including those survey months where a person was reported out of the labor force as part of the same spell. As the last column in Panel B of Table D7 shows, the resulting coefficient on unemployment duration remains positive and becomes even significant, suggesting that even for the SCE, our longitudinal results are not biased upward. Note that the nature of this exercise is somewhat different from the one further above, as the imputation procedure implies that we give more weight to individuals with more or longer transitions into out of the labor force.

The lack of updating over the spell seems pervasive across different groups of job seekers. When we regress the gradient of perceptions over the spell of unemployment, we find few characteristics that correlate significantly with it. For example, in the KM survey, measures of impatience, risk aversion or available savings do not correlate with the beliefs gradient.[1] We find a positive within-person correlation between liquidity constraints and the perceived probability - a job seeker reports a higher job-finding probability when liquidity constraints become binding - but controlling for liquidity constraints does not attenuate the positive impact of duration on beliefs.

To conclude, we have done extensive robustness checks and find that our results of the lack of negative updating of beliefs over the unemployment spell is a very robust result in both the SCE and the KM survey.

[1] Impatience in the KM survey is measured by the choice of a $20 incentive payment at the beginning of the survey over the option of a $40 incentive payment after the first 12 weeks of the survey. Risk aversion is elicited as the willingness to take risks on a scale from 0 to 10.

Table D7: Linear Regressions of Elicitations on Unemployment Duration, Robustness Checks (SCE)

| Dependent Variable (Unless Otherwise Stated in Footnote): Elicited 3-Month Job-Finding Probability | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Panel A.** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Unemployment Duration, in Months | 0.0022 | -0.0027 | 0.0034 | 0.0012 | -0.0004 | 0.0006 | -0.0001 | 0.0024 | 0.0017 |
| | (0.0064) | (0.0065) | (0.0068) | (0.0065) | (0.0043) | (0.0068) | (0.0003) | (0.0021) | (0.0016) |
| F.E. | S | S | S | S | S | S | S | I | I & T |
| Observations | 2,281 | 2,280 | 2,597 | 1,890 | 2,096 | 1,912 | 2,278 | 2,281 | 2,281 |
| $R^2$ | 0.824 | 0.835 | 0.789 | 0.868 | 0.790 | 0.813 | 0.823 | 0.806 | 0.827 |
| **Panel B.** | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
| Unemployment Duration, in Months | 0.0021 | 0.0033 | 0.0023 | 0.0022 | 0.0014 | 0.0011 | -0.0022 | 0.0057 | 0.0096*** |
| | (0.0063) | (0.0069) | (0.0064) | (0.0064) | (0.0094) | (0.0046) | (0.0053) | (0.0104) | (0.0032) |
| F.E. | S | S | S | S | S | S | S | S | S |
| Observations | 2,281 | 2,278 | 2,278 | 2,281 | 1,070 | 1,211 | 765 | 1,516 | 3,059 |
| $R^2$ | 0.824 | 0.824 | 0.824 | 0.824 | 0.876 | 0.806 | 0.812 | 0.825 | 0.672 |

*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers, ages 20-65, in the SCE. Column (1) reports the baseline results from Column 4 in Table 4; column (2) reports results where use the elicited 12-month job-finding probability as dependent variable; column (3) reports the results where we did not trim the sample for inconsistent answers between the two survey questions (i.e., where the 3-month probability was larger than the 12-month probability); in column (4) we excluded answers with a probability of 50 percent; in column (5) we excluded answers with a probability of 100 percent; in column (6) we excluded answers where the person was employed at the next interview; column (7) reports results with self-reported duration as the independent variable; column (8) reports results with individual fixed effects (I) instead of spell fixed effects (S); column (9) reports results where we control for individual fixed effects (I) and calendar time (T) fixed effects (month of survey); columns (10)-(13) control for the national unemployment rate, the state unemployment rate, the change in the state unemployment rate and quarterly GDP growth (in this order); column (14) reports results for those unemployed 6 months or less; column (15) reports results for those unemployed more than 6 months; column (16) reports results for spells with a transition into out of the labor force; column (17) reports results for continuous unemployment spells ending in a UE transition; column (18) reports results where missing elicitations in months out of the labor force are linearly inter- or extrapolated from months when unemployed. Robust standard errors (clustered at the spell level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

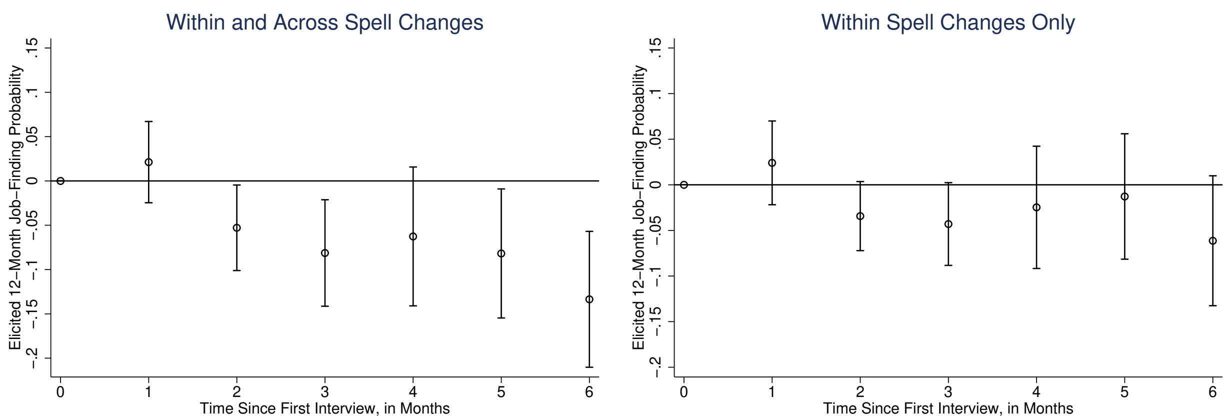
Table D8: Linear Regressions of Elicitations on Unemployment Duration, Robustness Checks (KM Survey)

| | Dependent Variable (Unless Otherwise Stated in Footnote): Elicited 1-Month Job-Finding Probability | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Unemployment Duration, | 0.022*** | 0.013 | 0.020*** | 0.021** | 0.022*** | 0.020** | 0.017** | 0.021*** | 0.030** |
| in Months | (0.008) | (0.008) | (0.008) | (0.009) | (0.008) | (0.008) | (0.008) | (0.008) | (0.014) |
| Person Fixed Effects | x | x | x | x | x | x | x | x | x |
| Observations | 4,435 | 4,269 | 4,486 | 3,862 | 4,292 | 4,105 | 4,044 | 4,272 | 4,435 |
| $R^2$ | 0.902 | 0.887 | 0.897 | 0.920 | 0.890 | 0.901 | 0.901 | 0.899 | 0.902 |

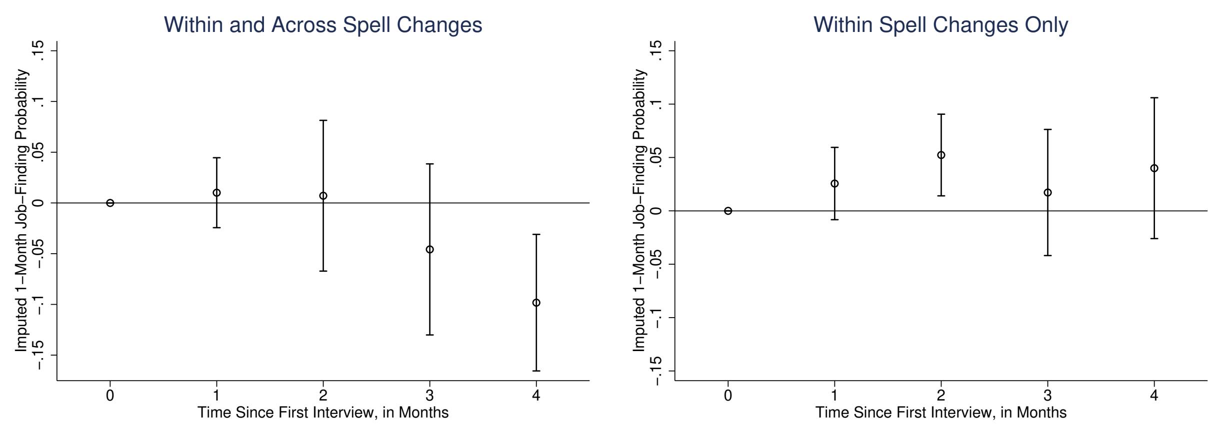
*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers, ages 20-65, in the KM survey. Column (1) reports the baseline results from Table 4; Column (2) uses the inverse of the expected duration question as dependent variable (see footnote 11 in the maintext for details); Column (3) reports the results where we did not trim the sample for inconsistent answers between the two survey questions (i.e., where the difference between the probability question and the inverse of the remaining duration was more than 75 percentage points apart); Column (4) reports results where we excluded answers with a probability of 50 percent; Column (5) reports results where we excluded probabilities of 80 percent or more; Column (6) reports the results where we excluded answers where the person reported in the following 4 weeks that she or he accepted a job or was working; Column (7) reports the results where we excluded answers where the respondent had previously received but not accepted a job offer; Column (8) reports results with self-reported duration as the independent variable; Column (9) reports results where we control for the monthly unemployment rate in New Jersey as reported by the BLS. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

Figure D6: Elicitations about Job Finding by Time since First Interview

**Panel A.** Elicited 12-Month Job-Finding Probability (SCE)



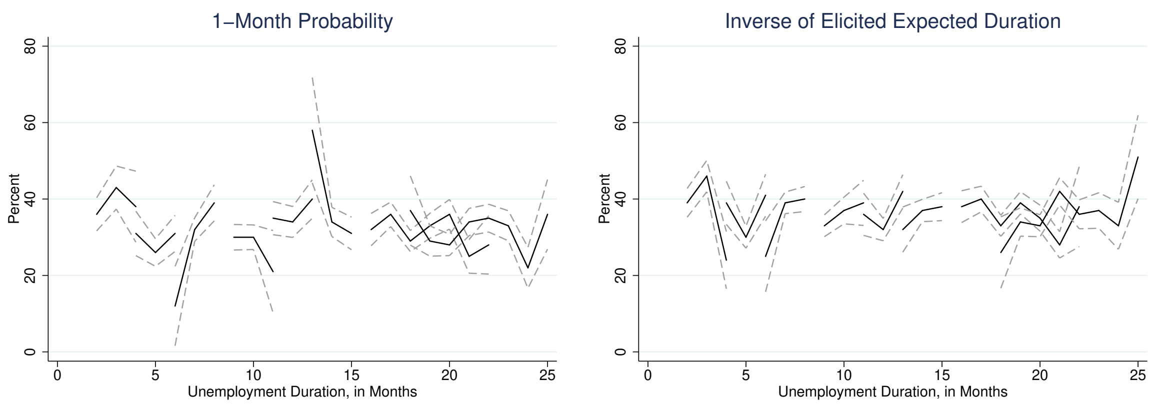

**Panel B.** Inverse of Elicited Expected Remaining Duration (KM Survey)



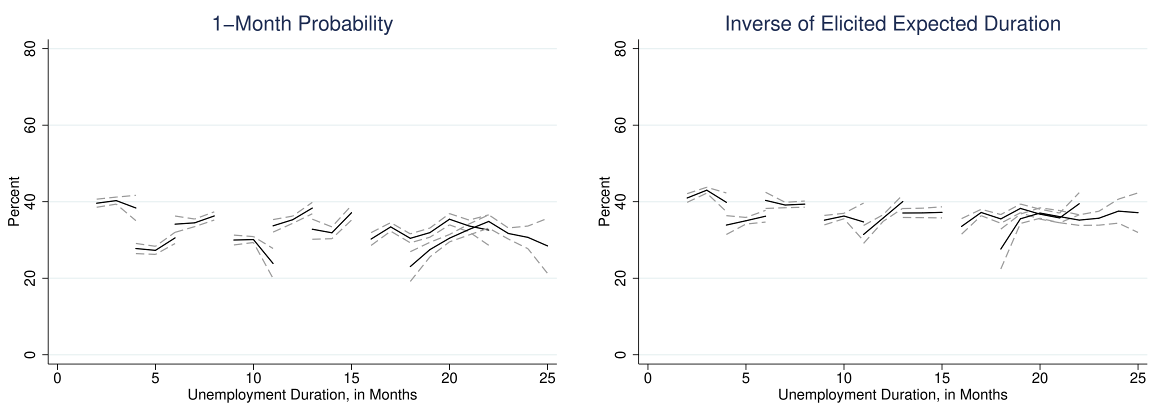

*Notes*: Survey weights are used for the averages shown in the figures, and the samples are restricted to unemployed workers, ages 20-65. The figures show the elicited job-finding probabilities by months since the first interview, in which a belief question was administered, for the SCE (Panel A) and the KM survey (Panel B). The left-hand side figures show the raw averages of the elicited job-finding probabilities, whereas the right-hand side figures remove individual fixed effects from the elicited job-finding probabilities. The elicited expected remaining duration of unemployment (in weeks) is inverted to make it comparable to a 1-month job-finding probability, computed as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration. The bars indicate the 95% confidence interval.

Figure D7: Average of Elicited Job-Finding Probabilities, by Duration of Unemployment and by Cohort (KM Survey)

**Panel A. Raw Averages**




**Panel B. Removing Individual Fixed Effects**




*Notes*: Survey weights are used for the averages shown in the figures, and the samples are restricted to unemployed workers in the KM survey, ages 20-65. The figure shows averages for each month of unemployment duration and by cohort. The figures in Panel A show raw averages, whereas the figures in Panel B show averages after removing individual fixed effects and adding the average for each cohort. The figures does not show averages for month-cohort bins with less than 10 observations. The elicited expected remaining duration of unemployment (in weeks) is inverted to make it comparable to a 1-month job-finding probability, computed as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration. The bars indicate the 95% confidence interval.

### D.3.3 Further Evidence on Elicited Beliefs

In this section, we provide some evidence on the relationship between eliced beliefs and job search behavior as reported in the KM survey. We also report some evidence on how job-seekers' beliefs respond to aggregate indicators of labor market tightness in the SCE.

Table D9 reports the results of regressions on reported search behavior in the KM data. We find that across job seekers, the self-reported reservation wage bears a negative association with the 1-month probability though statistically insignificant, whereas time spent on job search activities is a positive predictor of the elicited 1-month probability (significant at the 1 percent level). Overall, these results are, at least qualitatively, in line with what one could expect from a simple search model with a reservation wage choice and endogenous search effort: the reservation wage has a negative effect on the probability of accepting and thus finding a job, whereas search effort increases the probability of finding a job. The causality, however, may well run in the opposite direction. Job seekers who update positively on the probability of receiving a job offer are likely to increase their reservation wage. Indeed, we find some evidence for this in column 4: controlling for individual fixed effects, job seekers who decrease their reservation wage, reduce at the same time their expected remaining duration, though for the 1-month probability question the relationship remains small and insignificant. Reverse causality may confound the relationship between job finding beliefs and search effort as well. Controlling for individual fixed effects, the correlation between perceived job finding and search disappears.

When deciding how hard to search, the perceived returns to search are key as well (Spinnewijn [2015]). The survey gauges job seekers' perceived control by asking whether they could increase their job finding chances by spending more time searching for a job. Interestingly, the vast majority of job seekers state that they cannot. Table D9 shows, controlling for search effort, that workers who report a positive return to search at the margin also report higher job-finding probabilities.

We revisit the question on the role of beliefs for job search in our structural analysis in Section 5, where we specify a search model allowing for heterogeneous beliefs about the primitives of the job search environment and calibrate this model targeting the true and perceived job finding in our data. Note that our analysis in the statistical model in Section 4 abstracts from job search decisions and does not rely on any assumption about how beliefs affect job search either.

Table D10 shows workers' perceptions respond to aggregate indicators of job finding in the SCE. We find that for unemployed individuals there is no significant relationship between the national or state-level unemployment rate and the 3-month perception, though standard errors are relatively large. We do find, however, a highly significant positive correlation with the elicited probability that the stock market will rise and a highly significant negative correlation with the elicited probability that the unemployment rate will rise. This suggests that unemployed job seekers take into account their perceptions about aggregate conditions when expressing their perceptions about individual job finding (or vice versa), but their perceptions about aggregate conditions seem ill-informed. These results thus seem to suggest that unemployed workers' perceptions under-react to aggregate indicators of their employment prospects.

Table D9: Linear Regressions of Elicitations on Time Spent on Job Search and the Reservation Wage (KM Survey)

| Dependent variable: | Elicited 1-Month Job-Finding Probability | | Elicited Remaining Duration (Inverted) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dummy for Control Belief | 0.0884*** | -0.0109 | 0.1053*** | 0.0533* |
| | (0.0253) | (0.0230) | (0.0197) | (0.0307) |
| Time Spent on Job Search (Hours per Week) | 0.0013** | -0.0014 | 0.0011* | 0.0008 |
| | (0.0006) | (0.0011) | (0.0006) | (0.0014) |
| Log(Hourly Reservation Wage) | -0.0304 | -0.0109 | -0.0477 | 0.1346* |
| | (0.0346) | (0.0758) | (0.0298) | (0.0812) |
| Reservation Commuting Distance, in Minutes | -0.0002 | -0.0009 | -0.0008* | -0.0003 |
| | (0.0006) | (0.0013) | (0.0005) | (0.0013) |
| Controls | x | | x | |
| Individual F.E. | | x | | x |
| Observations | 3,967 | 4,059 | 3,905 | 3,984 |
| $R^2$ | 0.151 | 0.916 | 0.132 | 0.893 |

*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers, ages 20-65, in the KM survey. Expected remaining duration (inverted) is calculated as $1-(1-\frac{1}{x})^4$, where $x$ is the elicited expected remaining duration of unemployment (in weeks). The dummy for control belief is set to one for respondents who believe that chances of finding a job increase if they spent more time searching. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

Table D10: Linear Regressions of Macroeconomic Measures on Elicitations (SCE)

| Dependent Variable: Elicited 3-Month Job-Finding Probability | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| National Unemployment Rate | 0.812 | | | |
| | (1.369) | | | |
| National Job Openings Rate | 2.597 | | | |
| | (2.460) | | | |
| State Unemployment Rate | | -0.236 | -0.561 | |
| | | (0.832) | (0.763) | |
| Elicited Prob(Rise in US Unemployment) | | | | 0.165*** |
| | | | | (0.039) |
| Elicited Prob(Rise in Stock Prices) | | | | -0.079* |
| | | | | (0.042) |
| Demographics | x | x | x | x |
| State FE | | | x | x |
| Observations | 2,593 | 2,593 | 2,593 | 2,569 |
| $R^2$ | 0.103 | 0.102 | 0.175 | 0.187 |

*Notes*: Survey weights are used in all regressions. All samples are restricted to unemployed workers in the SCE, ages 20-65. Robust standard errors (clustered at the individual level) are in parentheses. Asteriks indicate statistical significance at the *0.1, **0.05 and ***0.01 level.

# E Statistical Framework

## E.1 Additional Details on Setup of Model, Distributional Assumptions and Functional Forms

We propose to parametrize our model relatively parsimoniously. Baseline job-finding rates, $T_i$, follow the Beta distribution with shape parameters $\alpha$ and $\beta$. The Beta distribution is defined over the interval $[0,1]$ and is quite flexible in terms of its shape. Note that for our exercise here it is important that there is a continuum of job-finding probabilities or at least a large number. Assuming two types for the job-finding probabilities and estimating their relative mass is not an attractive option, because our observed elicitations are reported on the interval between 0 and 1. A model with only two underlying job-finding rates thus would not perform well in matching the distribution of these elicitations.

The transitory component of the job-finding rate, $\tau_{i,d}$, follows a uniform distribution subject to the bounds $[-T_i, \frac{1}{(1-\theta)^d} - T_i]$, and with masspoint(s) at the bounds of this interval such that $E(\tau_{i,d}|T_i) = 0$ for all $T_i$. More precisely, $\tau|T_i$ follows a uniform distribution on the interval $[\max(-\sigma_\tau, -T_i), \min(\sigma_\tau, \frac{1}{(1-\theta)^d} - T_i)]$, with a masspoint at the bound of this interval if a bound is binding, such that $E(\tau_{i,d}|T_i) = 0$ for all $T_i$.

Random error in perceptions or elicitations, $\varepsilon_{i,d}$, follows a uniform distribution on the interval $[-\sigma_\varepsilon, \sigma_\varepsilon]$ subject to the bounds $[-b_0 - b_1\hat{T}^3_{i,d}, 1 - b_0 - b_1 T^3_{i,d}]$, and with masspoint(s) at the bounds of this interval such that $E(\varepsilon_{i,d}|T^3_{i,d}) = 0$ for all $T^3_{i,d}$. More precisely, $\varepsilon|\hat{T}^3_{i,d}$ follows a uniform distribution on the interval $[\max(-\sigma_\varepsilon, -b_0 - b_1 T^3_{i,d}), \min(\sigma_\varepsilon, 1 - b_0 - b_1 T^3_{i,d})]$, with a masspoint at the bound of this interval if a bound is binding, such that $E(\varepsilon_{i,d}|T^3_{i,d}) = 0$ for all $T^3_{i,d}$.

We assume that the maximum duration for each job seeker is two years, but we relax this assumptions in a set of robustness checks, where we allow for a maximum duration of up to five years.

As discussed before, the identification of heterogeneity does not rely on particular distribution functions for $T_i$, $\tau_{i,d}$ and $\varepsilon_{i,d}$, and we test the sensitivity of our results to alternative distributional assumptions. Finally, job-finding rates depreciate at a geometric rate over the unemployment spell in our baseline specification, with $\theta_d = (1-\theta)^d$. In an alternative specification, we assume a piece-wise linear specification for the depreciation where $\theta_d = 1 - d\theta$ if $d \leq 12$ and $\theta_d = 1 - 12\theta$ otherwise.

## E.2 Model Fit and Estimation Results for Restricted Models

In this section, we provide additional details regarding the estimation of the statistical model in the paper. Table E1 shows details on the model fit for the baseline model. Table E2 shows additional details on the restricted versions of the model discussed in the paper and shown in Table 6.

Table E1: Matched Moments

| Moment | Symbol | Value in Data | Value in Model |
|---|---|---|---|
| Average of Realized 3-Month Job-Finding Rates: | | | |
| ... at 0-3 Months of Unemployment | $m_{F_{03}}$ | 0.642 | 0.646 |
| ... at 4-6 Months of Unemployment | $m_{F_{46}}$ | 0.472 | 0.456 |
| ... at 7 Months of Unemployment or More | $m_{F_{7+}}$ | 0.256 | 0.262 |
| Average of Elicited 3-Month Job-Finding Probability (Deviation from Realized): | | | |
| ... at 0-3 Months of Unemployment | $m_{Z_{03}} - m_{F_{03}}$ | $-0.026$ | $-0.027$ |
| ... at 4-6 Months of Unemployment | $m_{Z_{46}} - m_{F_{46}}$ | 0.057 | 0.058 |
| ... at 7 Months of Unemployment or More | $m_{Z_{7+}} - m_{F_{7+}}$ | 0.153 | 0.147 |
| Average of Monthly Innovations in Elicitations | $m_{dZ}$ | 0.008 | 0.006 |
| Variance of 3-Month Elicitations | $s_Z^2$ | 0.093 | 0.093 |
| Covariance of 3-Month Elicitations and Job Finding | $c_{Z,F}$ | 0.054 | 0.056 |
| Covariance of 3-Month Elicitations and Job Finding in 3 Months | $c_{Z_d,F_{d+3}}$ | 0.025 | 0.024 |

*Notes*: Survey weights are used for all data moments, which are based on the SCE. The sample is restricted to unemployed workers, ages 20-65, and includes only interviews that were followed by three consecutive monthly interviews. The monthly innovations in elicitations refers to monthly individual-level changes in the elicited 3-month job-finding probability.

Table E2: Parameter Estimates and Model Fit for Restricted Versions of the Model

**Panel A. Parameter Estimates and Selected Moments**

| | (1) Baseline | (2) $\theta = 0$ | (3) No heterog. in $T_{i,d}$ | (4) $\sigma_\tau = 0$ | (5) $b_1 = 1$ | (6) $b_0 = 0$ $b_1 = 1$ |
|---|---|---|---|---|---|---|
| $E(T_i)$ | 0.403 | 0.405 | 0.318 | 0.419 | 0.294 | 0.332 |
| $Var(T_i)$ | 0.044 | 0.051 | 0.000 | 0.070 | 0.019 | 0.019 |
| $\sigma_\tau$ | 0.334 | -0.275 | 0.000 | 0.000 | 0.201 | 0.245 |
| $\theta$ | 0.017 | 0.000 | 0.106 | -0.044 | 0.012 | 0.015 |
| $b_0$ | 0.265 | 0.269 | 0.341 | 0.27 | 0.065 | 0.000 |
| $b_1$ | 0.55 | 0.538 | 0.349 | 0.543 | 1.000 | 1.000 |
| $\sigma_\varepsilon$ | 0.453 | 0.453 | 0.442 | 0.454 | 0.374 | 0.38 |
| $Var_0(T^3_{i,0})$ | 0.076 | 0.079 | 0.000 | 0.085 | 0.051 | 0.047 |
| $Var_0(T^3_i)$ | 0.057 | 0.065 | 0.000 | 0.085 | 0.040 | 0.033 |
| $Var_0(Z^3_{i,0})$ | 0.081 | 0.082 | 0.066 | 0.083 | 0.087 | 0.085 |
| TD | 0.459 | 0.454 | 0.430 | 0.452 | 0.301 | 0.315 |
| LD | 0.070 | 0.000 | 0.430 | -0.116 | 0.054 | 0.071 |
| $1 - LD/TD$ (in %) | 84.71 | 100.00 | 0.000 | 125.67 | 82.20 | 77.53 |

**Panel B. Model Fit**

| | Data | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|---|
| *Targeted Moments:* | | | | | | | |
| $m_{F_{03}}$ | 0.642 | 0.646 | 0.640 | 0.598 | 0.63 | 0.571 | 0.627 |
| $m_{F_{46}}$ | 0.472 | 0.456 | 0.446 | 0.460 | 0.418 | 0.464 | 0.520 |
| $m_{F_{7+}}$ | 0.256 | 0.262 | 0.263 | 0.218 | 0.262 | 0.321 | 0.373 |
| $m_{Z_{03}} - m_{F_{03}}$ | -0.026 | -0.027 | -0.027 | -0.049 | -0.018 | 0.064 | 0.000 |
| $m_{Z_{46}} - m_{F_{46}}$ | 0.057 | 0.058 | 0.062 | 0.041 | 0.078 | 0.066 | -0.001 |
| $m_{Z_{7+}} - m_{F_{7+}}$ | 0.153 | 0.147 | 0.148 | 0.199 | 0.151 | 0.066 | 0.00 |
| $m_{dZ}$ | 0.008 | 0.006 | 0.007 | -0.013 | 0.006 | 0.004 | 0.006 |
| $s^2_Z$ | 0.093 | 0.093 | 0.092 | 0.069 | 0.092 | 0.098 | 0.099 |
| $c_{Z,F}$ | 0.054 | 0.056 | 0.056 | 0.012 | 0.054 | 0.059 | 0.058 |
| $c_{Z_d,F_{d+3}}$ | 0.025 | 0.024 | 0.025 | 0.010 | 0.029 | 0.034 | 0.034 |
| Weighted SSR | | 0.470 | 0.695 | 49.198 | 2.514 | 12.251 | 21.613 |

*Notes*: The table shows the estimation results for restricted versions of the baseline model, where Panel A provides the parameter estimates and selected moments capturing the heterogeneity in true and perceived job finding and the decline in true job finding over the first 12 months of the unemployment spell and Panel B shows the model fit. The total decline (TD) is the difference in the average of the true 3-month job-finding probability between job seekers at beginning of the unemployment spell and those still unemployed after 12 months, $E_0(T^3_{i,0}) - E_{12}(T^3_{i,12})$. The longitudinal decline (LD) is the individual-level decline in the true 3-month job-finding probability over the first 12 months of the unemployment spell, averaged across all job seekers, $E_0(T^3_{i,0} - T^3_{i,12})$. $1 - \frac{LD}{TD}$ is the share of the total decline that the model attributes to heterogeneity in true job-finding probabilities and the resulting process of dynamic selection.

### E.3 Robustness

In this section, we study the robustness of our results when using alternative specifications, alternative moments, and alternative functional forms and distributions. Tables E3 and E4 show results for various robustness checks discussed briefly in the paper.

First, we assess the estimated heterogeneity in our statistical model using the heterogeneity that is not predictable using observable characteristics. The robustness check we perform is to estimate the model on a set of residualized moments, i.e., the moments obtained from the residuals of a set of linear regressions of the 3-month belief question and of the 3-month job-finding rate on the same set of demographic controls as in Table 2. The estimation results are shown in Table E4 and the moments in Table E5 . Overall, the estimation results are very similar to the baseline, with the role played by true duration dependence being again close to zero. Of course, the extent of ex-ante heterogeneity is estimated to be smaller in this robustness check, as the effects of observables are parsed out from all moments. We also obtain a comparable estimate for the slope coefficient $b_1$ of 0.559, which suggests that the relationship between observed heterogeneity in job finding and beliefs is similar as the relationship between unobserved heterogeneity in job finding and beliefs.

Second, we probe the robustness of our findings to alternative assumptions about the functional form and distributions as well as extensions of the model, as reported in Tables E3 and E4. Without discussing these estimates in detail, the table shows that the parameter estimates are very stable across all of the results reported in the table. In particular, our results are robust to assuming that $T_i$ follows the Gamma distribution (2), and to assuming that $\varepsilon$ follows a bounded normal distribution, which no longer satisfies mean-independence of the error term (3). Our results are also robust to assuming piecewise linear true duration dependence instead of geometric depreciation (4), extending the horizon of the model to 5 years (5), and doing both (6). We also extend the model to allow for completely persistent elicitation errors (i.e., $\varepsilon_{i,d} = \varepsilon_i$) and find that it has no impact on our estimation results (7). This is also true when we extend the model to allow for bunching at 0, 0.5 and 1 of the elicited beliefs, by imposing on the baseline model that *any* belief in the intervals $(0, 0.1]$, $[0.4, 0.6]$ resp. $[0.9, 1)$ are reset to the bunching points 0, 0.5 resp. 1. Despite these relatively strong assumptions about the nature of bunching, the results of the estimation appear not to be affected (8). This suggests that the variations in elicitations across (rather than within) these intervals is the dominant source of variation that is relevant for identification of the key parameters in the model. We also report the results for a model (9), where a share $\alpha$ of individuals has random elicitations ($Z^3_{i,d} = b_0 + \varepsilon$) and a share $1 - \alpha$ correctly perceives their job finding prospects ($Z^3_{i,d} = T^3_{i,d}$). The model results are very similar to our baseline, and the value of $1 - \alpha$ is close to the value of $b_1$, suggesting that $b_1$ in our baseline may instead capture the share of individuals who perceive their job finding prospects correctly. Our results are also very similar when using the residualized data moments as discussed before (10), or when excluding individuals with recall expectations when generating the data moments (11). Furthermore, the results do not change either when restricting the set of moments by using only 0-6 and 7+ months for the time intervals and dropping the mean of monthly innovations, so that the model is exactly identified (12), or when using the inverse of the bootstrapped variances on the diagonal of the weighting matrix (and zero otherwise) instead of the full variance-covariance matrix as the weighting matrix (13).

Finally, in (14) we show the results for the extended model, where $\hat{\theta} \neq \theta$, as discussed in detail in the paper. The targeted moments and the model fit are reported in Table E6. Figures E1 and E2 show the estimated duration dependence in true and perceived job finding for both the baseline model and the extended model. Overall, the duration dependence in both true job finding and the bias looks very similar across the two models. For this reason, not surprisingly, the fit of the restricted model in column 15 of Table E4 where $\hat{\theta} = \theta$ is close to the fit of the unrestricted version of the extended model in column 14.

To summarize, we find that our results for the baseline model are very robust to alternative assumptions about functional form and distributional assumptions as well as to extensions of our baseline model.

Table E3: Robustness Checks

| **Panel A. Parameter Estimates and Selected Moments** | | (1) Baseline | (2) Gamma $T_i$ | (3) Normal $\varepsilon$ | (4) Linear Depreciation | (5) Horizon=5y | (6) Horizon=5y Lin. Depr. | (7) Persistent errors | (8) Bunching | (9) Share $\alpha$ w/ rand. elicit. |
|---|---|---|---|---|---|---|---|---|---|---|
| $E(T_i)$ | | 0.403 | 0.401 | 0.405 | 0.406 | 0.380 | 0.383 | 0.401 | 0.403 | 0.402 |
| $Var(T_i)$ | | 0.045 | 0.044 | 0.044 | 0.046 | 0.038 | 0.041 | 0.041 | 0.044 | 0.043 |
| $\sigma_\tau$ | | 0.334 | 0.335 | 0.358 | 0.334 | 0.321 | 0.299 | 0.378 | 0.350 | 0.340 |
| $\theta$ | | 0.017 | 0.018 | 0.02 | 0.017 | 0.007 | 0.003 | 0.025 | 0.019 | 0.018 |
| $b_0$ | | 0.265 | 0.263 | 0.236 | 0.265 | 0.269 | 0.269 | 0.259 | 0.274 | 0.589 |
| $b_1$ | | 0.550 | 0.553 | 0.619 | 0.549 | 0.542 | 0.540 | 0.556 | 0.529 | 0.558* |
| $\sigma_\varepsilon$ | | 0.453 | 0.452 | 0.287 | 0.453 | 0.455 | 0.455 | 0.466 | 0.440 | 0.490 |
| $Var_0(T^3_{i,0})$ | | 0.076 | 0.076 | 0.078 | 0.077 | 0.072 | 0.073 | 0.077 | 0.077 | 0.076 |
| $Var_0(T^3_i)$ | | 0.057 | 0.056 | 0.056 | 0.058 | 0.053 | 0.056 | 0.053 | 0.056 | 0.056 |
| $Var_0(Z^3_{i,0})$ | | 0.081 | 0.081 | 0.08 | 0.081 | 0.081 | 0.082 | 0.084 | 0.081 | 0.079 |
| TD | | 0.459 | 0.459 | 0.462 | 0.464 | 0.414 | 0.418 | 0.460 | 0.461 | 0.458 |
| LD | | 0.070 | 0.075 | 0.083 | 0.071 | 0.030 | 0.013 | 0.107 | 0.079 | 0.077 |
| $100 - 100LD/TD$ | | 84.71 | 83.56 | 82.15 | 84.82 | 92.66 | 96.79 | 76.73 | 82.87 | 83.17 |
| **Panel B. Model Fit** | Data | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| *Targeted Moments:* | | | | | | | | | | |
| $m_{F_{03}}$ | 0.642 | 0.646 | 0.646 | 0.648 | 0.646 | 0.643 | 0.642 | 0.647 | 0.646 | 0.647 |
| $m_{F_{46}}$ | 0.472 | 0.456 | 0.457 | 0.457 | 0.452 | 0.478 | 0.473 | 0.459 | 0.456 | 0.458 |
| $m_{F_{7+}}$ | 0.256 | 0.262 | 0.262 | 0.261 | 0.263 | 0.255 | 0.256 | 0.260 | 0.261 | 0.264 |
| $m_{Z_{03}} - m_{F_{03}}$ | -0.026 | -0.027 | -0.026 | -0.027 | -0.027 | -0.025 | -0.025 | -0.028 | -0.027 | -0.025 |
| $m_{Z_{46}} - m_{F_{46}}$ | 0.057 | 0.058 | 0.058 | 0.058 | 0.060 | 0.050 | 0.051 | 0.057 | 0.059 | 0.057 |
| $m_{Z_{7+}} - m_{F_{7+}}$ | 0.153 | 0.147 | 0.146 | 0.148 | 0.147 | 0.153 | 0.153 | 0.148 | 0.147 | 0.144 |
| $m_{dZ}$ | 0.008 | 0.006 | 0.006 | 0.006 | 0.007 | 0.008 | 0.008 | 0.006 | 0.007 | 0.006 |
| $s^2_Z$ | 0.093 | 0.093 | 0.093 | 0.093 | 0.093 | 0.093 | 0.092 | 0.092 | 0.093 | 0.093 |
| $c_{Z,F}$ | 0.054 | 0.056 | 0.056 | 0.056 | 0.056 | 0.055 | 0.055 | 0.056 | 0.056 | 0.056 |
| $c_{Z_d,F_{d+3}}$ | 0.025 | 0.024 | 0.024 | 0.024 | 0.024 | 0.024 | 0.025 | 0.023 | 0.024 | 0.024 |
| Weighted SSR | | 0.4698 | 0.4598 | 0.4660 | 0.6032 | 0.0289 | 0.0427 | 0.4320 | 0.3923 | 0.5059 |

*Notes*: The table shows the estimation results of robustness checks of the statistical model, where Panel A provides the parameter estimates and selected moments capturing the heterogeneity in true and perceived job finding and the decline in true job finding over the first 12 months of the unemployment spell and Panel B shows the model fit. The total decline (TD) is the difference in the average of the true 3-month job-finding probability between job seekers at beginning of the unemployment spell and those still unemployed after 12 months, $E_0(T^3_{i,0}) - E_{12}(T^3_{i,12})$. The longitudinal decline (LD) is the individual-level decline in the true 3-month job-finding probability over the first 12 months of the unemployment spell, averaged across all job seekers, $E_0(T^3_{i,0} - T^3_{i,12})$. $1 - \frac{LD}{TD}$ is the share of the total decline that the model attributes to heterogeneity in true job-finding probabilities and the resulting process of dynamic selection. *In this model, a share $\alpha$ of individuals has random elicitations ($Z^3_{i,d} = b_0 + \varepsilon$) and a share $1 - \alpha$ correctly perceives their job finding prospects ($Z^3_{i,d} = T^3_{i,d}$). We report the estimated value of $1 - \alpha$ instead of $b_1$.

Table E4: Robustness Checks (Different Targeted Moments)

| Parameter Estimates and Selected Moments: | (1) Baseline | (10) Residualized Moments | (11) Excluding Recall | (12) Exact Identification | (13) Diagonal W | (14) Extended Model | (15) Extended w/ $\hat{\theta} = \theta$ |
|---|---|---|---|---|---|---|---|
| $E(T_i)$ | 0.403 | 0.348 | 0.403 | 0.404 | 0.403 | 0.421 | 0.412 |
| $Var(T_i)$ | 0.045 | 0.022 | 0.043 | 0.048 | 0.045 | 0.048 | 0.042 |
| $\sigma_\tau$ | 0.334 | 0.279 | 0.360 | 0.279 | 0.341 | 0.477 | 0.484 |
| $\theta$ | 0.017 | 0.011 | 0.022 | 0.012 | 0.016 | 0.022 | 0.035 |
| $\hat{\theta}$ | — | — | — | — | — | 0.037 | 0.035 |
| $b_0$ | 0.265 | 0.258 | 0.265 | 0.269 | 0.273 | 0.291 | 0.270 |
| $b_1$ | 0.550 | 0.559 | 0.548 | 0.543 | 0.538 | 0.518 | 0.542 |
| $\sigma_\varepsilon$ | 0.453 | 0.425 | 0.453 | 0.456 | 0.456 | 0.459 | 0.455 |
| $Var_0(T^3_{i,0})$ | 0.076 | 0.053 | 0.076 | 0.076 | 0.077 | 0.091 | 0.087 |
| $Var_0(T^3_i)$ | 0.057 | 0.036 | 0.054 | 0.062 | 0.057 | 0.057 | 0.051 |
| $Var_0(Z^3_{i,0})$ | 0.081 | 0.071 | 0.081 | 0.081 | 0.081 | 0.083 | 0.083 |
| TD | 0.459 | 0.332 | 0.462 | 0.462 | 0.460 | 0.482 | 0.480 |
| LD | 0.070 | 0.048 | 0.095 | 0.046 | 0.068 | 0.089 | 0.151 |
| $1 - LD/TD$ (in %) | 84.71 | 85.5 | 79.49 | 90.08 | 85.25 | 81.64 | 68.46 |
| Weighted SSR | 0.470 | 0.260 | 0.487 | 0.000 | 0.204 | 3.999 | 4.141 |

*Notes*: The table shows the estimation results for robustness checks of the statistical model. In this table, the Weighted SSR refers to different sets of moments, so they cannot be compared across specifications in this table. See Table E5 for the residualized moments targeted in the estimation of the results reported in column 4. The total decline (TD) is the difference in the average of the true 3-month job-finding probability between job seekers at beginning of the unemployment spell and those still unemployed after 12 months, $E_0(T^3_{i,0}) - E_{12}(T^3_{i,12})$. The longitudinal decline (LD) is the individual-level decline in the true 3-month job-finding probability over the first 12 months of the unemployment spell, averaged across all job seekers, $E_0(T^3_{i,0} - T^3_{i,12})$. $1 - \frac{LD}{TD}$ is the share of the total decline that the model attributes to heterogeneity in true job-finding probabilities and the resulting process of dynamic selection.

Table E5: Matched Moments (Residualized)

| Moment | Symbol | Value in Data | Value in Model |
|---|---|---|---|
| Average of Realized 3-Month Job-Finding Rates: | | | |
| ... at 0-3 Months of Unemployment | $m_{F_{03}}$ | 0.642 | 0.641 |
| ... at 4-6 Months of Unemployment | $m_{F_{46}}$ | 0.513 | 0.525 |
| ... at 7 Months of Unemployment or More | $m_{F_{7+}}$ | 0.369 | 0.373 |
| Average of Elicited 3-Month Job-Finding Probability (Deviation from Realized): | | | |
| ... at 0-3 Months of Unemployment | $m_{Z_{03}} - m_{F_{03}}$ | $-0.026$ | $-0.025$ |
| ... at 4-6 Months of Unemployment | $m_{Z_{46}} - m_{F_{46}}$ | 0.042 | 0.026 |
| ... at 7 Months of Unemployment or More | $m_{Z_{7+}} - m_{F_{7+}}$ | 0.098 | 0.094 |
| Average of Monthly Innovations in Elicitations | $m_{dZ}$ | 0.009 | 0.006 |
| Variance of 3-Month Elicitations | $s_Z^2$ | 0.078 | 0.078 |
| Covariance of 3-Month Elicitations and Job Finding | $c_{Z,F}$ | 0.036 | 0.038 |
| Covariance of 3-Month Elicitations and Job Finding in 3 Months | $c_{Z_d,F_{d+3}}$ | 0.020 | 0.020 |

*Notes*: Survey weights are used for all data moments, which are based on the SCE. The sample is restricted to unemployed workers, ages 20-65, and includes only interviews that were followed by three consecutive monthly interviews. Moments are computed based on residuals from a regression on dummies for gender, race, ethnicity, household income, educational attainment, and age and age squared. Note that the raw mean of the variables in the full sample is added to the residual. The monthly innovations in elicitations refers to monthly individual-level changes in the elicited 3-month job-finding probability.

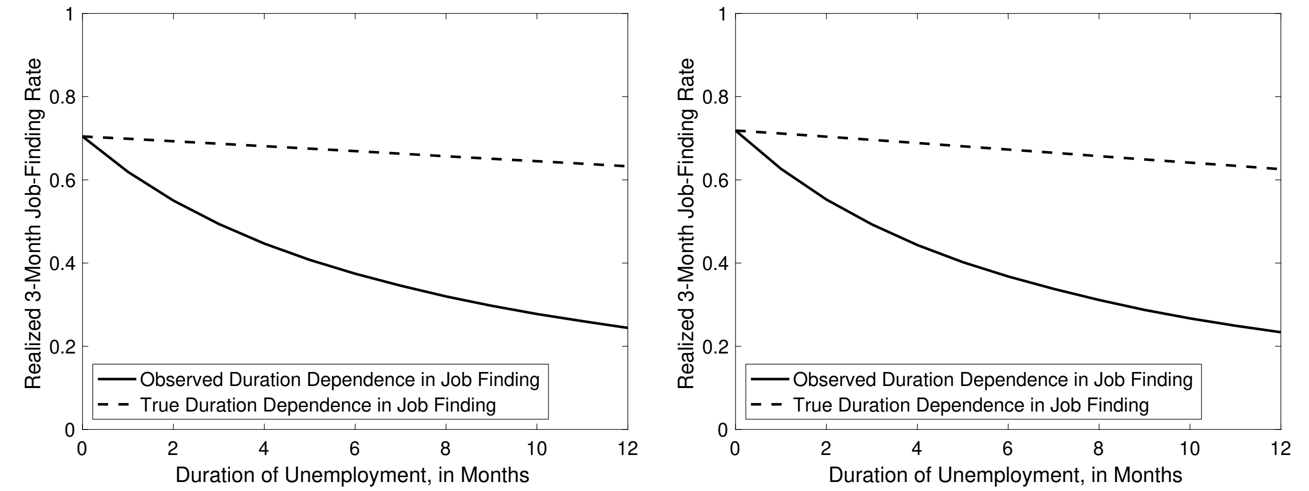
Table E6: Matched Moments (Extended Model)

| Moment | Symbol | Value in Data | Value in Model |
|---|---|---|---|
| Average of Realized 3-Month Job-Finding Rates: | | | |
| ... at 0-3 Months of Unemployment | $m_{F_{03}}$ | 0.642 | 0.656 |
| ... at 4-6 Months of Unemployment | $m_{F_{46}}$ | 0.472 | 0.452 |
| ... at 7 Months of Unemployment or More | $m_{F_{7+}}$ | 0.256 | 0.250 |
| Average of Elicited 3-Month Job-Finding Probability (Deviation from Realized): | | | |
| ... at 0-3 Months of Unemployment | $m_{Z_{03}} - m_{F_{03}}$ | $-0.026$ | $-0.030$ |
| ... at 4-6 Months of Unemployment | $m_{Z_{46}} - m_{F_{46}}$ | 0.057 | 0.060 |
| ... at 7 Months of Unemployment or More | $m_{Z_{7+}} - m_{F_{7+}}$ | 0.153 | 0.153 |
| Average of Monthly Innovations in Elicitations | $m_{dZ}$ | 0.008 | 0.009 |
| Variance of 3-Month Elicitations: | | | |
| ... at 0-6 Months of Unemployment | $s^2_{Z_{06}}$ | 0.098 | 0.091 |
| ... at 7 Months of Unemployment or More | $s^2_{Z_{7+}}$ | 0.073 | 0.079 |
| Covariance of 3-Month Elicitations and Job Finding: | | | |
| ... at 0-6 Months of Unemployment | $c_{Z_{06},F_{06}}$ | 0.056 | 0.055 |
| ... at 7 Months of Unemployment or More | $c_{Z_{7+},F_{7+}}$ | 0.025 | 0.029 |
| Covariance of 3-Month Elicitations and Job Finding in 3 Months | $c_{Z_d,F_{d+3}}$ | 0.025 | 0.021 |

*Notes*: Survey weights are used for all data moments, which are based on the SCE. The sample is restricted to unemployed workers, ages 20-65, and includes only interviews that were followed by three consecutive monthly interviews. The monthly innovations in elicitations refers to monthly individual-level changes in the elicited 3-month job-finding probability.

Figure E1: Duration Dependence in Job Finding in Baseline and Extended Model

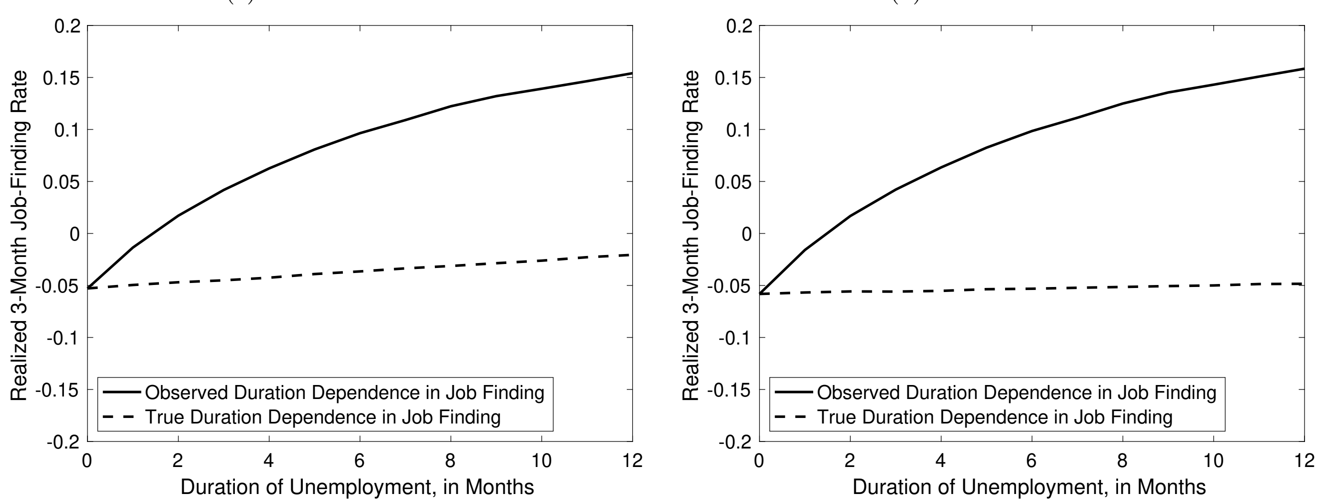
(a) Baseline Model (b) Extended Model




*Notes*: The figure shows further estimation results of the baseline model (a) and the extended model (b). The observed duration dependence in job finding shows the averages of the realized 3-month job-finding rate at duration $d$, averaged across job seekers still unemployed after $d$ months of unemployment, $E_d(T^3_{i,d})$. The true duration dependence in job finding shows the realized 3-month job-finding rate at duration $d$, averaged across all job seekers, $E_0(T^3_{i,d})$.

Figure E2: Duration Dependence in Biases in Perceptions in Baseline and Extended Model

(a) Baseline Model (b) Extended Model




*Notes*: The figure shows further estimation results of the baseline model (a) and the extended model (b). The observed duration dependence in bias shows the differences in the averages of the perceived and realized 3-month job-finding rate at duration $d$, averaged across job seekers still unemployed after $d$ months of unemployment, $E_d(Z^3_{i,d}) - E_d(T^3_{i,d})$. The true duration dependence in bias shows the differences in the averages of the perceived and realized 3-month job-finding rate at duration $d$, averaged across all job seekers, $E_0(Z^3_{i,d}) - E_0(T^3_{i,d})$.

### E.4 Identification

In this section, we provide further details on the identification of the parameters in the statistical model. We proceed in two steps:

First, we prove that in a two-period version of the statistical model, where $\sigma_\tau = 0$, all other parameters are a function of moments with an empirical counterpart in the data and thus are identified.

Second, we provide a formal identification argument in the two-period model where $\sigma_\tau > 0$, and then show that in the full model a monotone relationship exists between $\sigma_\tau$ and the moment $c_{Z_d,F_{d+3}}$, conditional on having identified all other parameters of the model.

Third, we extend the two-period model to allow for $\hat{\theta} \neq \theta$ and prove that in the version of this model, where $\sigma_\tau = 0$, all other parameters are a function of moments with an empirical counterpart in the data and thus are identified. We then show in the full model that a monotonic relationship exists between $\hat{\theta}$ and the moments $c_{F_{7+},Z_{7+}}$ and $s^2_{Z_{7+}}$, conditional on having identified all other parameters of the model.

#### E.4.1 Identification in two-period model with $\sigma_\tau = 0$

**Proposition 2.** *In a two-period version of the statistical model with measurement error, $var(\varepsilon)$, that is independent of $T_{i,d}$ and with $\sigma_\tau = 0$, the parameters $b_0$, $b_1$, and $\theta$ as well as the mean and the variance of the persistent component of job-finding rates, $E(T_i)$ and $var(T_i)$, and the variance of the elicitation error, $var(\varepsilon)$, are identified by the moment conditions for: (1) the means of the elicited job-finding probabilities in period 1 and 2, $m_{Z_1}$ and $m_{Z_2}$, (2) the means of the realized job-finding rates in period 1 and 2, $m_{F_1}$ and $m_{F_2}$, (3) the covariance of realized job finding and elicited job-finding probabilities in period 1, $c_{F_1,Z_1}$, and (4) the variance of elicited job-finding probabilities in period 1, $s^2_{Z_1}$.*

*Proof.* We start by assuming that there are only two periods, and that $\sigma_\tau = 0$. In this case, we can write down the moment conditions for the moments mentioned in the proposition above as:

$$\begin{aligned}
m_{Z_1} &= b_0 + b_1 E_1(T_i) &(21)\\
m_{Z_2} &= b_0 + b_1(1-\theta)E_2(T_i) &(22)\\
m_{F_1} &= E_1(T_i) &(23)\\
m_{F_2} &= (1-\theta)E_2(T_i) &(24)\\
c_{F_1,Z_1} &= cov_1(F_{i,1}, b_1 T_i) &(25)\\
s^2_{Z_1} &= b_1^2 var_1(T_i) + var(\varepsilon) &(26)
\end{aligned}$$

where sub-indices 1 and 2 on the moments stands for the sample of survivors. Note that $E_1(T_i) = E(T_i)$ and $var_1(T_i) = var(T_i)$, i.e. the moments for the sample of survivors in period 1 correspond to the population moments. The first two moments directly pin down $b_0$ and $b_1$:

$$\begin{aligned}
b_1 &= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} &(27)\\
b_0 &= m_{Z_1} - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_1} &(28)
\end{aligned}$$

Then, we can write:

$$
\begin{aligned}
c_{F_1,Z_1} &= cov_1(F_{i,1}, b_1 T_i) \\
&= b_1[E_1(F_{i,1}T_i) - E_1(F_{i,1})E_1(T_i)] \\
&= b_1[E_1(E_1(F_{i,1}T_i|T_i)) - E_1(T_i)^2] \\
&= b_1[E_1(T_i^2) - E(T_i)^2] \\
&= b_1 var_1(T_i) = b_1 var(T_i) \qquad (29)
\end{aligned}
$$

Hence, we can pin down the mean and the variance of $T_i$ from moment conditions (23) and (29):

$$
\begin{aligned}
E(T_i) &= m_{F_1} \qquad (30) \\
var(T_i) &= \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} \qquad (31)
\end{aligned}
$$

We next note that we can re-write the expected value of $T_i$, conditional on survival to period 2 as:

$$
\begin{aligned}
E_2(T_i) &= \frac{E_1\big[T_i(1-T_i)\big]}{1-E(T_i)} \\
&= \frac{E_1(T_i) - E_1(T_i^2)}{1-E_1(T_i)} \\
&= \frac{E_1(T_i)(1-E_1(T_i)) - var_1(T_i)}{1-E_1(T_i)} \qquad (32)
\end{aligned}
$$

Substituting this into the moment condition for $m_{F_2}$, we get:

$$
m_{F_2} = (1-\theta)\frac{E_1(T_i)(1-E_1(T_i)) - var_1(T_i)}{1-E(T_i)} \qquad (33)
$$

Rearranging and using equation (30), we get:

$$
\begin{aligned}
\theta &= 1 - \frac{m_{F_2}(1-m_{F_1})}{m_{F_1}(1-m_{F_1}) - var_1(T_i)} \\
&= 1 - \frac{m_{F_2}(1-m_{F_1})}{m_{F_1}(1-m_{F_1}) - \frac{m_{F_2}-m_{F_1}}{m_{Z_2}-m_{Z_1}} c_{F_1,Z_1}} \\
&= 1 - \frac{(m_{Z_2}-m_{Z_1})(m_{F_2}(1-m_{F_1}))}{m_{F_1}(1-m_{F_1}) - (m_{F_2}-m_{F_1})c_{F_1,Z_1}} \qquad (34)
\end{aligned}
$$

Finally, given $b_1$, we can solve for $var(\varepsilon)$ by using the moment condition for $s^2_{Z_1}$:

$$
var(\varepsilon) = s^2_{Z_1} - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} c_{F_1,Z_1} \qquad (35)
$$

Since $var(\varepsilon)$ is increasing in $\sigma_\varepsilon$, the equation implies a value for $\sigma_\varepsilon$.

In conclusion, equations 27, 28, 30, 31, 34 and 35 solve parameters $b_0$, $b_1$, $\theta$ and moments $E(T_i)$, $var(T_i)$ and $var(\varepsilon)$ for *any* distribution of these variables as function of moments that we observe in the data ($m_{Z_1}$, $m_{Z_2}$, $m_{F_1}$, $m_{F_2}$, $c_{F_1,Z_1}$ and $s^2_{Z_1}$). The two-period model with $\sigma_\tau = 0$ is thus identified.

□

**Proposition 3.** *In a two-period version of the statistical model with both a measurement error, $\varepsilon$, that is independent of $T_{i,d}$, as well as a non-classical measurement error of the form $\eta = c_0 + c_1 T_{i,d}$, and with $\sigma_\tau = 0$, the parameters $\tilde{b}_0 = b_0$, $\tilde{b}_1 = b_1 + c_1$, and $\theta$ as well as the mean and the variance of the persistent component of job-finding rates, $E(T_i)$ and $var(T_i)$, and the variance of the classical elicitation error, $var(\varepsilon)$, are identified by the moment conditions for: (1) the means of the elicited job-finding probabilities in period 1 and 2, $m_{Z_1}$ and $m_{Z_2}$, (2) the means of the realized job-finding rates in period 1 and 2, $m_{F_1}$ and $m_{F_2}$, (3) the covariance of realized job finding and elicited job-finding probabilities in period 1, $c_{F_1,Z_1}$, and (4) the variance of elicited job-finding probabilities in period 1, $s^2_{Z_1}$.*

*Proof.* As argued in the main text of the paper, the model continues to be identified exactly in the presence of non-classical measurement error, as long as it is a linear in $T_i$. The proof is almost trivial, as we can re-express equations 21, 22, 25 and 26 from above as:

$$
\begin{aligned}
m_{Z_1} &= (b_0 + c_0) + (b_1 + c_1) E_1(T_i) && (36)\\
m_{Z_2} &= (b_0 + c_0) + (b_1 + c_1)(1 - \theta) E_2(T_i) && (37)\\
c_{F_1,Z_1} &= cov_1(F_i, (b_1 + c_1) T_i) && (38)\\
s^2_{Z_1} &= (b_1 + c_1)^2 var_1(T_i) + var(\varepsilon) && (39)
\end{aligned}
$$

or:

$$
\begin{aligned}
m_{Z_1} &= \tilde{b}_0 + \tilde{b}_1 E_1(T_i) && (40)\\
m_{Z_2} &= \tilde{b}_0 + \tilde{b}_1 (1 - \theta) E_2(T_i) && (41)\\
c_{F_1,Z_1} &= cov_1(F_i, \tilde{b}_1 T_i) && (42)\\
s^2_{Z_1} &= (\tilde{b}_1)^2 var_1(T_i) + var(\varepsilon) && (43)
\end{aligned}
$$

These moment conditions are identical to the ones in the model without non-classical measurement error, except that we replaced $b_0$ and $b_1$ with $\tilde{b}_0$ and $\tilde{b}_1$. It follows from the proof for Proposition 2 that the parameters $\tilde{b}_0$, $\tilde{b}_1$, $\theta$ and moments $E(T_i)$, $var(T_i)$ and $var(\varepsilon)$ are functions of moments that we observe in the data ($m_{Z_1}$, $m_{Z_2}$, $m_{F_1}$, $m_{F_2}$, $c_{F_1,Z_1}$ and $s^2_{Z_1}$). The two-period model with $\sigma_\tau = 0$ and non-classical measurement error of the form $\eta = c_0 + c_1 T_i$ is thus identified. □

### E.4.2 Identification of $\sigma_\tau$

Our conjecture is that in a two-period version of the statistical model with measurement error, $\varepsilon$, that is independent of $T_{i,d}$, with transitory shocks to job finding, $\tau_{i,d}$, that are independent of $T_i$, and with $G(T_i)$ following a two-parameter distribution, the parameters $b_0$, $b_1$, $\theta$, and $\sigma_\tau$ as well as the mean and the variance of the persistent component of job-finding rates, $E(T_i)$ and $var(T_i)$, and the variance of the elicitation error, $var(\varepsilon)$, are identified by the moment conditions for: (1) the means of the elicited job-finding probabilities in period 1 and 2, $m_{Z_1}$ and $m_{Z_2}$, (2) the means of the realized job-finding rates in period 1 and 2, $m_{F_1}$ and $m_{F_2}$, (3) the covariance of realized job finding and the elicited job-finding probabilities in period 1, $c_{F_1,Z_1}$, (4) the covariance of realized job finding in period 2 and the elicited job-finding probabilities in period 1, $c_{F_2,Z_1}$, and (5) the variance of the elicited job-finding probabilities in period 1, $s^2_{Z_1}$.

We again consider a model with only two periods, period 1 and 2. In this case, we can write down the moment conditions for the moments mentioned in the proposition above as:

$$\begin{aligned}
m_{Z_1} &= b_0 + b_1 E_1(T_i + \tau_{i,1}) && (44)\\
m_{Z_2} &= b_0 + b_1(1-\theta)E_2(T_i + \tau_{i,2}) && (45)\\
m_{F_1} &= E_1(T_i + \tau_{i,1}) && (46)\\
m_{F_2} &= (1-\theta)E_2(T_i + \tau_{i,2}) && (47)\\
c_{F_1,Z_1} &= cov_1(F_i, b_1(T_i + \tau_{i,1})) && (48)\\
c_{F_2,Z_1} &= cov_2(F_i, b_1(T_i + \tau_{i,1})) && (49)\\
s^2_{Z_1} &= b_1^2 var_1(T_i + \tau_{i,1}) + var(\varepsilon) && (50)
\end{aligned}$$

The first two moments again directly pin down $b_0$ and $b_1$:

$$\begin{aligned}
b_1 &= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} && (51)\\
b_0 &= m_{Z_1} - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_1} && (52)
\end{aligned}$$

We can again re-write the expectation conditional on survival to period 2, now of $T_i + \tau_{i,1}$, as:

$$\begin{aligned}
E_2(T_i + \tau_{i,1}) &= \frac{E_1\left[(T_i + \tau_{i,1})(1 - T_i - \tau_{i,1})\right]}{1 - E_1(T_i + \tau_{i,1})}\\
&= \frac{E_1(T_i) - E_1(T_i^2) - E_1(\tau_{i,1}^2)}{1 - E_1(T_i)}\\
&= \frac{E_1(T_i)(1 - E_1(T_i)) - var_1(T_i) - var_1(\tau_{i,1})}{1 - E_1(T_i)} && (53)
\end{aligned}$$

because $E_1(\tau_{i,1}) = E_1(T_i \tau_{i,1}) = 0$. Similarly, we obtain

$$E_2(T_i + \tau_{i,2}) = \frac{E_1(T_i)(1 - E_1(T_i)) - var_1(T_i)}{1 - E_1(T_i)} \tag{54}$$

because $E_1(\tau_{i,1}) = E_1(T_i\tau_{i,2}) = E_1(T_i\tau_{i,1}) = E_1(\tau_{i,1}\tau_{i,2}) = 0$. Hence, we can re-write:

$$\begin{aligned}
c_{F_1,Z_1} &= cov_1(F_{i,1}, b_1(T_i + \tau_{i,1})) \\
&= b_1[E_1(F_{i,1}(T_i + \tau_{i,1})) - E_1(F_i)E_1(T_i + \tau_{i,1})] \\
&= b_1[E_1(E_1(F_{i,1}(T_i + \tau_{i,1})|T_i, \tau_{i,1})) - E_1(T_i + \tau_{i,1})E_1(T_i + \tau_{i,1})] \\
&= b_1[E_1(E_1((T_i + \tau_{i,1})(T_i + \tau_{i,1})|T_i, \tau_{i,1})) - E_1(T_i + \tau_{i,1})E_1(T_i + \tau_{i,1})] \\
&= b_1[E_1((T_i + \tau_{i,1})^2) - E_1(T_i + \tau_{i,1})E_1(T_i + \tau_{i,1})] \\
&= b_1[E_1((T_i^2 + 2T_i\tau_{i,1} + \tau_{i,1}^2) - E_1(T_i)E_1(T_i)] \\
&= b_1[E_1((T_i^2) + E_1(\tau_{i,1}^2) - E_1(T_i)E_1(T_i)] \\
&= b_1[var_1(T_i) + var_1(\tau_{i,1})] \qquad (55)
\end{aligned}$$

because $E_1(T_i\tau_{i,1}) = 0$. Similarly, we obtain:

$$\begin{aligned}
c_{F_2,Z_1} &= cov_2(F_{i,2}, b_1(T_i + \tau_{i,1})) \\
&= b_1[E_2(F_{i,2}(T_i + \tau_{i,1})) - E_2(F_{i,2})E_2(T_i + \tau_{i,1})] \\
&= b_1[E_2(E_2(F_{i,2}(T_i + \tau_{i,1})|T_i, \tau_{i,1})) - (1-\theta)E_2(T_i + \tau_{i,2})E_2(T_i + \tau_{i,1})] \\
&= b_1[E_2((1-\theta)(T_i + \tau_{i,2})(T_i + \tau_{i,1})) - (1-\theta)E_2(T_i + \tau_{i,2})E_2(T_i + \tau_{i,1})] \\
&= b_1(1-\theta)[E_2((T_i + \tau_{i,2})(T_i + \tau_{i,1})) - E_2(T_i + \tau_{i,2})E_2(T_i + \tau_{i,1})] \\
&= b_1(1-\theta)\left[\frac{E_1(T_i^2) - E_1(T_i^3) - E_1(T_i\tau_{i,1}^2)}{1 - E_1(T_i)} - \Big(E_1(T_i) - \frac{var_1(T_i)}{1 - E_1(T_i)}\Big)\Big(E_1(T_i) - \frac{var_1(T_i) + var_1(\tau_{i,1})}{1 - E_1(T_i)}\Big)\right]
\end{aligned}$$

where the last equality uses the same steps as before to re-write the conditional expectation. Re-arranging terms and using $m_{F_1} = E_1(T_i)$, $m_{F_2} = (1-\theta)\Big[m_{F_1} - \frac{var_1(T_i)}{1-m_{F_1}}\Big]$ and $b_1 = \frac{m_{Z_2}-m_{Z_1}}{m_{F_2}-m_{F_1}}$, we get:

$$\begin{aligned}
c_{F_2,Z_1} &= b_1(1-\theta)\left[\frac{var_1(T_i) + m_{F_1}^2 - E_1(T_i^3) - E_1(T_i\tau_{i,1}^2)}{1 - m_{F_1}}\right] - b_1 m_{F_2}\Big(m_{F_1} - \frac{1}{b_1}\frac{c_{F_1,Z_1}}{1 - m_{F_1}}\Big) \\
&= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} \frac{m_{F_2}(1 - m_{F_1})}{m_{F_1}(1 - m_{F_1}) - var(T_i)} \left[\frac{var_1(T_i) + m_{F_1}^2 - E_1(T_i^3) - E_1(T_i\tau_{i,1}^2)}{1 - m_{F_1}}\right] \\
&\quad - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_2} m_{F_1} + m_{F_2}\frac{c_{F_1,Z_1}}{1 - m_{F_1}} \\
&= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} \frac{m_{F_2}(1 - m_{F_1})}{m_{F_1}(1 - m_{F_1}) - var(T_i)} \left[\frac{var_1(T_i) + m_{F_1}^2 - E_1(T_i^3)}{1 - m_{F_1}}\right] \\
&\quad - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_2} m_{F_1} + m_{F_2}\frac{c_{F_1,Z_1}}{1 - m_{F_1}}
\end{aligned}$$

Using equation (55) to get $var_1(T_i) = \frac{c_{F_1,Z_1}}{b_1} - var_1(\tau_{i,1})$, we can rearrange the equation above, to get:

$$\begin{aligned} c_{F_2,Z_1} &= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} \frac{m_{F_2}(1 - m_{F_1})}{m_{F_1}(1 - m_{F_1}) - \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} + var_1(\tau_{i,1})} \\ & \left[ \frac{\frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} - var_1(\tau_{i,1}) + m_{F_1}^2 - E_1(T_i^3)}{1 - m_{F_1}} \right] \\ & - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_2} m_{F_1} + m_{F_2} \frac{c_{F_1,Z_1}}{1 - m_{F_1}} \end{aligned}$$

For two-parameter distributions of $T_i$ where $E_1(T_i^3)$ is either implicitly or explicitly defined by the first two moments of the distribution, we can define a function $h(.,.)$, such that $E_1(T_i^3) = h(E_1(T_i), var_1(T_i))$, and thus:

$$\begin{aligned} c_{F_2,Z_1} &= \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} \frac{m_{F_2}(1 - m_{F_1})}{m_{F_1}(1 - m_{F_1}) - \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} + var_1(\tau_{i,1})} \\ & \left[ \frac{\frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} - var_1(\tau_{i,1}) + m_{F_1}^2 - h(m_{F_1}, \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} - var_1(\tau_{i,1}))}{1 - m_{F_1}} \right] \\ & - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} m_{F_2} m_{F_1} + m_{F_2} \frac{c_{F_1,Z_1}}{1 - m_{F_1}} \end{aligned} \tag{56}$$

While it is not possible to solve explicitly for $\sigma_\tau$, we note that for $h_2 \leq 0$, the right-hand side of the equation (56) above depends negatively on $\sigma_\tau$, and thus a solution for $\sigma_\tau$ exists.[2] A solution also exists for $h_2 \leq \tilde{h}$, where $\tilde{h}$ is some positive number, as long as $\tilde{h}$ is smaller than some upper bound $\bar{h}$.

Having solved for $var_1(\tau_{i,1})$, when a solution to equation (56) exists, we can then find a solution for the mean and variance of $T_i$:

$$E_1(T_i) = m_{F_1} \tag{57}$$

$$var_1(T_i) = \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} - var_1(\tau_{i,1}) \tag{58}$$

Rearranging and using equation (47), we also get:

$$\theta = 1 - \frac{m_{F_2}(1 - m_{F_1})}{m_{F_1}(1 - m_{F_1}) - \frac{m_{F_2} - m_{F_1}}{m_{Z_2} - m_{Z_1}} c_{F_1,Z_1} + var_1(\tau_{i,1})} \tag{59}$$

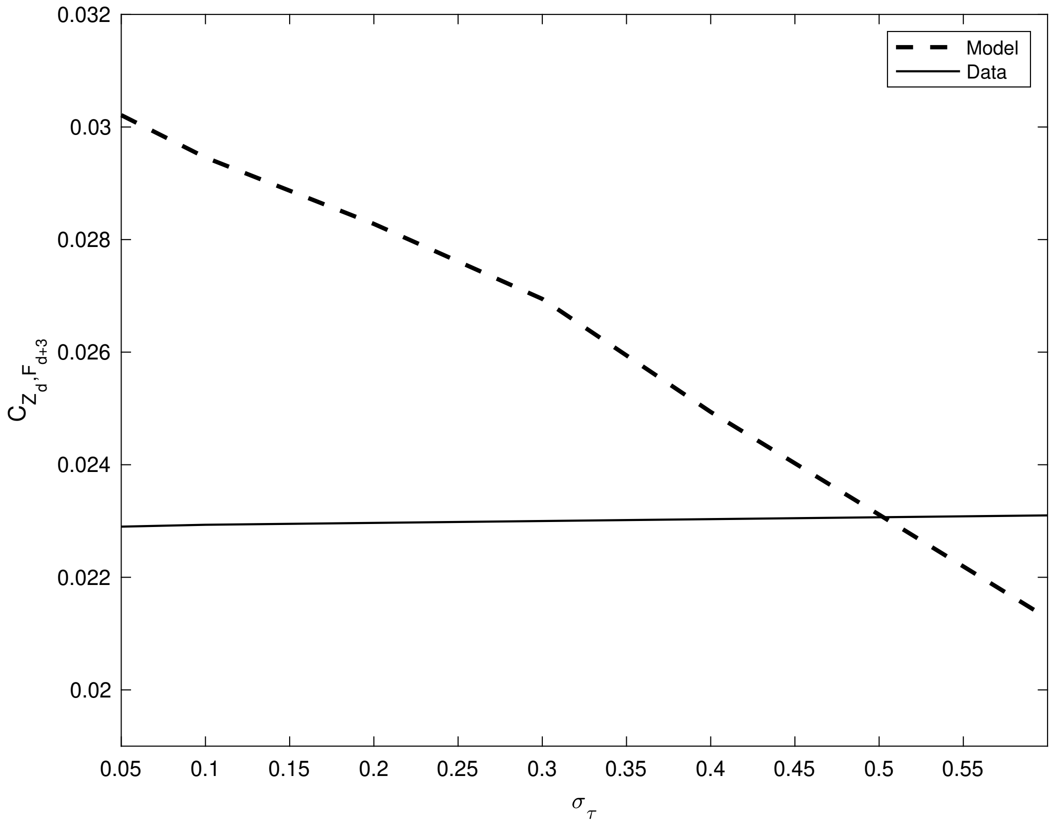
As before, given $b_1$, we can also solve for $var(\varepsilon)$ by using the moment condition for $s_{Z_1}^2$:

$$var(\varepsilon) = s_{Z_1}^2 - \frac{m_{Z_2} - m_{Z_1}}{m_{F_2} - m_{F_1}} c_{F_1,Z_1} \tag{60}$$

In conclusion, if a solution exists to equation (56), implicitly defining $\sigma_\tau$, we can solve for parameters $b_0$, $b_1$, $\sigma_\varepsilon$, and $\theta$ as well as the mean and variance of the persistent component of job-finding rates, $E(T_i) = E_1(T_i)$ and $var(T_i) = var_1(T_i)$, as a function of the moments $m_{Z_1}$, $m_{Z_2}$, $m_{F_1}$, $m_{F_2}$, $c_{F_1,Z_1}$,

[2] Note $\sigma_\tau$ is monotonically increasing in but not equal to $var_1(\tau_{i,1})$, because of the boundary conditions.

Figure E3: The relationship between $\sigma_\tau$ and the moment $c_{Z_d,F_{d+3}}$ in the estimated sub-model




$c_{F_2,Z_1}$, $s^2_{Z_1}$, as shown in equations 51, 52, 57, 58, 59 and 60.

To provide further evidence on identification of the parameter $\sigma_\tau$, we now proceed by showing that in the context of our estimated model (i.e., with more than two periods), there is a monotone mapping between the parameter $\sigma_\tau$ and the moment $C_{Z_d,F_{d+3}}$. More precisely, we estimate a sub-model of the baseline version of our statistical model for different levels of $\sigma_\tau$, by targeting all of the same moments *except* $C_{Z_d,F_{d+3}}$. Figure E3 shows that there is a monotone relationship between the level of $\sigma_\tau$ and the covariance of the elicited job-finding probabilities and the 3-month forward realized job-finding rates in this estimated sub-model, which shows that our parameter $\sigma_\tau$ is identified by the moment $C_{Z_d,F_{d+3}}$ in the full (baseline) model.

### E.4.3 Identification in extended model with $\hat{\theta} \neq \theta$

**Proposition 4.** *In a two-period version of the extended version of the statistical model with $\hat{\theta} \neq \theta$ and with measurement error, $\varepsilon$, that is independent of $T_{i,d}$, but with $\sigma_\tau = 0$, the parameters $b_0$, $b_1$, $\theta$, $\hat{\theta}$ as well as the mean and the variance of the persistent component of job-finding rates, $E(T_i)$ and $var(T_i)$, and the variance of the elicitation error, $var(\varepsilon)$, are identified by the moment conditions for: (1) the means of the elicited job-finding probabilities in period 1 and 2, $m_{Z_1}$ and $m_{Z_2}$, (2) the means of the realized job-finding rates in period 1 and 2, $m_{F_1}$ and $m_{F_2}$, (3) the covariance of realized job finding and the elicited job-finding probabilities in period 1, $c_{F_1,Z_1}$, (4) the variance of the elicited job-finding* *probabilities in period 1, $s^2_{Z_1}$, and (5) a statistic that depends on all these moments as well as the variance of the elicited job-finding probabilities in period 2, $s^2_{Z_2}$, and the covariance with realized job finding in period 2, $c_{F_2,Z_2}$.*

*Proof.* We assume that there are only two periods, and that $\sigma_\tau = 0$. In this case, we can write down the moment conditions for the moments mentioned in the proposition above as:

$$m_{Z_1} = b_0 + b_1 E_1(T_i) \tag{61}$$
$$m_{Z_2} = b_0 + b_1(1-\hat{\theta})E_2(T_i) \tag{62}$$
$$m_{F_1} = E_1(T_i) \tag{63}$$
$$m_{F_2} = (1-\theta)E_2(T_i) \tag{64}$$
$$c_{F_1,Z_1} = cov_1(F_{i,1}, b_1 T_i) \tag{65}$$
$$c_{F_2,Z_2} = cov_2(F_{i,2}, b_1 T_i) \tag{66}$$
$$s^2_{Z_1} = b_1^2 var_1(T_i) + var(\varepsilon) \tag{67}$$
$$s^2_{Z_2} = b_1^2(1-\hat{\theta})^2 var_2(T_i) + var(\varepsilon) \tag{68}$$

where sub-indices 1 and 2 on the moments stands for the sample of survivors. Note that $E_1(T_i) = E(T_i)$ and $var_1(T_i) = var(T_i)$, i.e. the moments for the sample of survivors in period 1 correspond to the population moments. One can express the additional moment condition (66) as follows:

$$\begin{aligned}
c_{F_2,Z_2} &= cov_2(F_{i,2}, b_1(1-\hat{\theta})T_i) \\
&= b_1(1-\hat{\theta})[E_2(F_{i,2}T_i) - E_2(F_{i,2})E(T_i)] \\
&= b_1(1-\hat{\theta})[E_2(E_2(F_{i,2}T_i|T_i)) - (1-\theta)E_2(T_i)^2] \\
&= b_1(1-\hat{\theta})(1-\theta)[E_2(T_i^2) - E_2(T_i)^2] \\
&= b_1(1-\hat{\theta})(1-\theta)var_2(T_i)
\end{aligned}$$

Re-arranging the moment conditions 61-67 and using equations (29) and (32), we thus get:

$$m_{Z_1} = b_0 + b_1 m_{F_1} \tag{69}$$
$$m_{Z_2} = b_0 + b_1 \frac{1-\hat{\theta}}{1-\theta} m_{F_2} \tag{70}$$
$$m_{F_1} = E_1(T_i) \tag{71}$$
$$m_{F_2} = (1-\theta)\left[m_{F_1} - \frac{var_1(T_i)}{1-m_{F_1}}\right] \tag{72}$$
$$c_{F_1,Z_1} = b_1 var_1(T_i) \tag{73}$$
$$c_{F_2,Z_2} = b_1(1-\hat{\theta})(1-\theta)var_2(T_i) \tag{74}$$
$$s^2_{Z_1} = b_1^2 var_1(T_i) + var(\varepsilon) \tag{75}$$
$$s^2_{Z_2} = b_1^2(1-\hat{\theta})^2 var_2(T_i) + var(\varepsilon) \tag{76}$$

The mean of the job-finding rate, $E(T_i) = E_1(T_i)$, is directly identified by moment condition in equation

(71).We then take the difference of the first two moment conditions:

$$m_{Z_1} - m_{Z_2} = b_1\big(m_{F_1} - \frac{1-\hat{\theta}}{1-\theta} m_{F_2}\big) \tag{77}$$

which gives $b_1$ as a function of moments, $\theta$ and $\hat{\theta}$. Next combine equations 73 and 75 and equations 74 and 76, to get:

$$s^2_{Z_1} = b_1 c_{F_1,Z_1} + var(\varepsilon) \tag{78}$$

$$s^2_{Z_2} = b_1 \frac{1-\hat{\theta}}{1-\theta} c_{F_2,Z_2} + var(\varepsilon) \tag{79}$$

and taking the difference, we get:

$$s^2_{Z_1} - s^2_{Z_2} = b_1(c_{F_1,Z_1} - \frac{1-\hat{\theta}}{1-\theta} c_{F_2,Z_2}) \tag{80}$$

Taking the ratio of equation 77 and 80, we get:

$$\frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}} = \frac{m_{F_1} - \frac{1-\hat{\theta}}{1-\theta} m_{F_2}}{c_{F_1,Z_1} - \frac{1-\hat{\theta}}{1-\theta} c_{F_2,Z_2}} \tag{81}$$

Rearranging:

$$\frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}}(c_{F_1,Z_1} - \frac{1-\hat{\theta}}{1-\theta} c_{F_2,Z_2}) = m_{F_1} - \frac{1-\hat{\theta}}{1-\theta} m_{F_2} \tag{82}$$

Rearranging further:

$$\frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}} c_{F_1,Z_1} - m_{F_1} = -\frac{1-\hat{\theta}}{1-\theta} m_{F_2} + \frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}} \frac{1-\hat{\theta}}{1-\theta} c_{F_2,Z_2} \tag{83}$$

Rearranging further:

$$\frac{1-\hat{\theta}}{1-\theta} = \frac{\frac{m_{Z_1}-m_{Z_2}}{s^2_{Z_1}-s^2_{Z_2}} c_{F_1,Z_1} - m_{F_1}}{\frac{m_{Z_1}-m_{Z_2}}{s^2_{Z_1}-s^2_{Z_2}} c_{F_2,Z_2} - m_{F_2}} \tag{84}$$

Equation 84 defines the ratio of $\frac{1-\hat{\theta}}{1-\theta}$ as a function of moments only. Using the ratio, one can rearrange equation 77 to get $b_1$ as a function of moments only:

$$b_1 = \frac{m_{Z_1} - m_{Z_2}}{m_{F_1} - \frac{1-\hat{\theta}}{1-\theta} m_{F_2}} \tag{85}$$

Using $b_1$, one can use equation 61 to get $b_0$ as a function of moments only:

$$b_0 \quad = \quad m_{Z_1} - b_1 m_{F_2} \tag{86}$$

Using $b_1$, one can use equation 73 to get $var(T_i) = var_1(T_i)$ as a function of moments only:

$$var(T_i) \quad = \quad \frac{c_{F_1,Z_1}}{b_1} \tag{87}$$

Using $b_1$ and $var(T_i)$, one can use equation 75 to get $var(\varepsilon)$ as a function of moments only:

$$var(\varepsilon) \quad = \quad s^2_{Z_1} - b_1^2 var(T_i) \tag{88}$$

Using $var(T_i)$, one can use equation 64 to get $\theta$ as a function of moments only:

$$\theta \quad = \quad \frac{(m_{F_1} - m_{F_2})(1 - m_{F_1}) - var(T_i)}{m_{F_1}(1 - m_{F_1}) - var(T_i)} \tag{89}$$

Using $\theta$ and equation 84, we get $\hat{\theta}$ as a function of moments only:

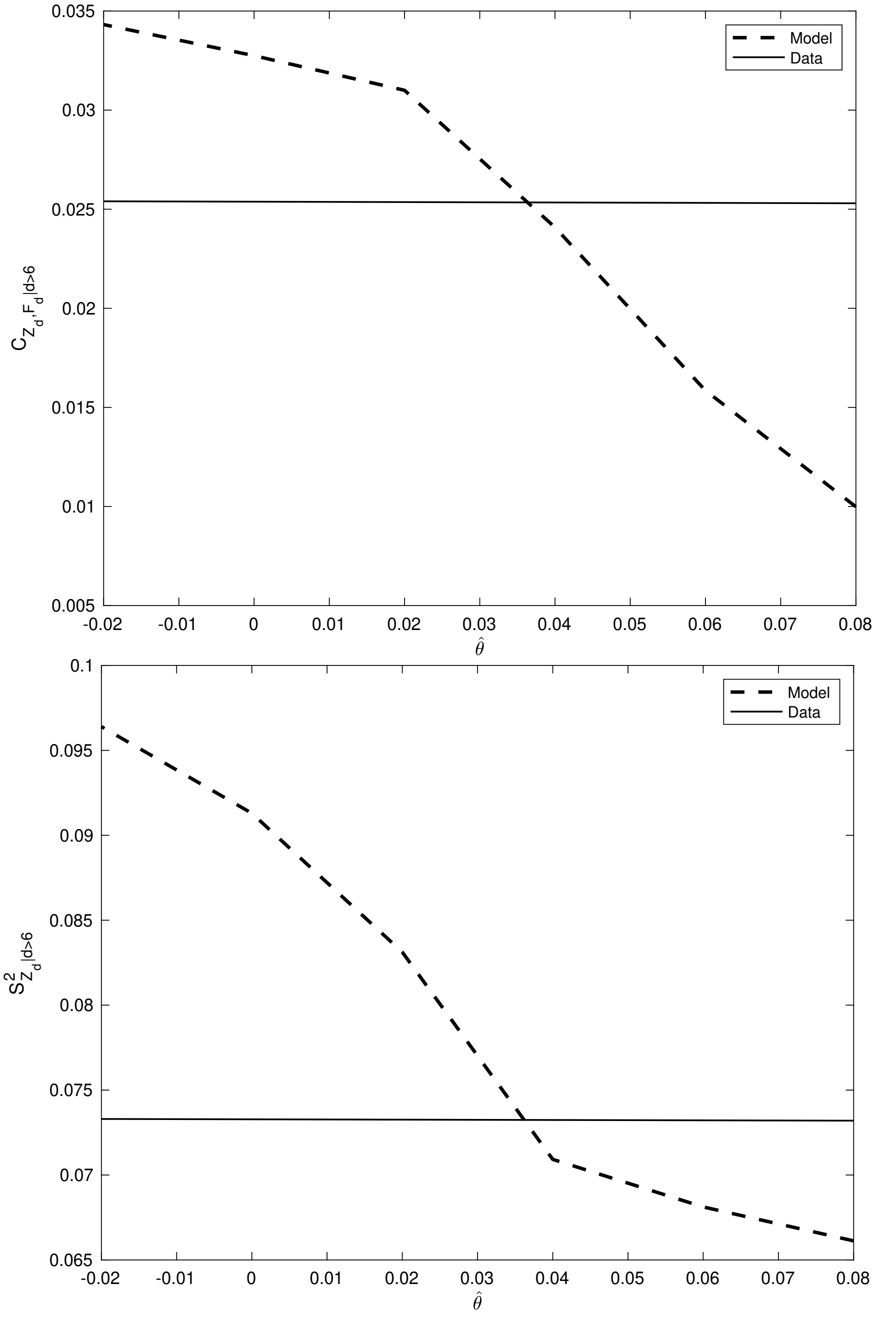
$$\hat{\theta} \quad = \quad 1 - (1-\theta)\frac{\frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}} c_{F_1,Z_1} - m_{F_1}}{\frac{m_{Z_1} - m_{Z_2}}{s^2_{Z_1} - s^2_{Z_2}} c_{F_2,Z_2} - m_{F_2}} \tag{90}$$

In conclusion, equations 85 (together with equation 84), 86, 63, 87, 88, 89 and 90 solve parameters $b_0$, $b_1$, $\theta$, $\hat{\theta}$ and population moments $E(T_i)$, $var(T_i)$ and $var(\varepsilon)$ for *any* distribution of these variables as function of moments that we observe in the data ($m_{Z_1}$, $m_{Z_2}$, $m_{F_1}$, $m_{F_2}$, $c_{F_1,Z_1}$, $s^2_{Z_1}$ and the ratio in equation 84). The two-period model with $\sigma_\tau = 0$ and $\hat{\theta} \neq \theta$ is thus identified.

$\square$

To provide evidence on the identification of the parameter $\hat{\theta}$ in the full version of the model that we estimate with the data, we proceed by showing that in the context of our estimated model (i.e., with more than two periods and $\sigma_\tau > 0$), there is a monotone mapping between $\hat{\theta}$ and the covariance in duration interval 7+, $c_{F_{7+},Z_{7+}}$ as well as a monotone mapping between $\hat{\theta}$ and the variance of elicitations in duration interval 7+, $s^2_{Z_{7+}}$. More precisely, we estimate a sub-model of the *extended* version of our statistical model, where $\hat{\theta} \neq \theta$, for different levels of $\hat{\theta}$, targeting all the same moments *except* $c_{F_{7+},Z_{7+}}$ and $s^2_{Z_{7+}}$. Figure E4 shows that there is a monotone relationship between the level of $\hat{\theta}$ and the covariance of the elicited job-finding probabilities and realized job finding in duration interval 7+ as well as a monotone relationship between level of $\hat{\theta}$ and the variance of elicitations in duration interval 7+. This shows that both of these moments provide variation that identifies the parameter $\hat{\theta}$ in the full (extended) model.

Figure E4: The relationship between $\hat{\theta}$ and moments $c_{F_{7+},Z_{7+}}$ and $s^2_{Z_{7+}}$ in the estimated sub-model



# F Structural Model

This section provides further details on the derivations and Proposition in the theoretical analysis and the calibration and counterfactual analysis in the numerical analysis in Section 5.

## F.1 Theoretical Analysis

In a stationary setting (with $\theta = 0$), the continuation values when unemployed and employed are:

$$\begin{aligned} U =& u(b_u) + \frac{1}{1+\delta} \max_R \left\{ U + \hat{\lambda} \int_R [V(w) - U] dF(w) ] \right\}, \\ V(w) =& u(w) + \frac{1}{1+\delta} \left\{ (1-\sigma) \left( V(w) + \lambda^e \int_w (V(x) - V(w)) dF(w) \right) + \sigma U \right\}, \end{aligned}$$

where $\sigma$ denotes the separation rate and $\lambda^e$ the arrival rate of job offers when employed. For an individual $i$, we have

$$T_{i,d} = \lambda_i [1 - F(R_i)].$$

Now dropping subindices, we can consider the impact on the job-finding rate $T$ of infinitesimal changes in $\lambda$ and $\hat{\lambda}$,

$$dT = [1 - F(R)] d\lambda - \lambda f(R) \frac{dR}{d\hat{\lambda}} d\hat{\lambda},$$

A change in $\lambda$ does not trigger a change in the reservation wage $R$ since it is only the perceived arrival rate that informs the agent's reservation wage. Rearranging this equation we get,

$$\frac{dT}{d\lambda} \frac{\lambda}{T} = 1 - \lambda \frac{f(R)}{1 - F(R)} \frac{dR}{d\hat{\lambda}} \frac{d\hat{\lambda}}{d\lambda}.$$

To unpack the $\frac{dR}{d\hat{\lambda}}$ term we consider the determination of the reservation wage. The reservation wage is defined by $U = V(R)$. Assuming $\sigma = \lambda^e = 0$, we can write,

$$V(R) = \frac{1+\delta}{\delta} u(R)$$

and thus

$$\frac{1+\delta}{\delta} u(R) = u + \frac{1}{1+\delta} \max_R \left\{ \frac{1+\delta}{\delta} u(R) + \hat{\lambda} \int_R \left[ V(w) - \frac{1+\delta}{\delta} u(R) \right] dF(w) \right\}.$$

We can totally differentiate this condition with respect to $R$ and $\hat{\lambda}$, applying the envelope theorem to the right hand side (i.e., $dU/dR = 0$) and assuming no job separation risk and no on-the-job search such that $V(w) = (1+\delta)u(w)/\delta$,

$$\frac{u'(R)}{1 - \frac{1}{1+\delta}} dR = \frac{1}{1+\delta} \left\{ \int_R \left[ \frac{u(w)}{1 - \frac{1}{1+\delta}} - \frac{u(R)}{1 - \frac{1}{1+\delta}} \right] dF(w) \right\} d\hat{\lambda}.$$

So, we can conclude

$$\frac{dR}{d\hat{\lambda}} = \frac{1}{1+\delta}\left\{ \int_R \left[\frac{u(w)-u(R)}{u'(R)}\right] dF(w)\right\}.$$

Combining this with our earlier result, we find

$$\begin{aligned}
\frac{dT}{d\lambda}\frac{\lambda}{T} &= 1 - \frac{1}{1+\delta}\lambda f(R)\frac{\int_R \left[\frac{u(w)-u(R)}{u'(R)}\right] dF(w)}{1-F(R)}\frac{d\hat{\lambda}}{d\lambda}, \\
&= 1 - \frac{1}{1+\delta}T\frac{f(R)}{1-F(R)}E\left[\frac{u(w)-u(R)}{u'(R)}\middle| w \geq R\right]\frac{d\hat{\lambda}}{d\lambda} \\
&\equiv 1 - \kappa\frac{d\hat{\lambda}}{d\lambda}.
\end{aligned}$$

This corresponds to equation (17) in the main text.

### F.1.1 Proof of Proposition 1

First, we consider the introduction of heterogeneity. That is, we assume $\lambda^j = \bar{\lambda} + d\lambda^j$ for $j = h, l$ with $d\lambda^h = -d\lambda^l \approx 0$, but we keep $\theta = 0$. We also assume an equal share of high and low types, $\phi = 1/2$. Now for small differences in actual and perceived arrival rates, we can approximate

$$\begin{aligned}
T_i &\approx T + \frac{dT}{d\lambda_i}d\lambda_i + \frac{dT}{d\hat{\lambda}_i}d\hat{\lambda}_i, \\
&= \bar{\lambda}[1-F(R)] + [1-F(R)]d\lambda_i - \bar{\lambda}f(R)\frac{dR}{d\hat{\lambda}_i}d\hat{\lambda}_i \\
&= \bar{\lambda}[1-F(R)] + [1-F(R)]\left[d\lambda_i - \kappa d\hat{\lambda}_i\right].
\end{aligned}$$

Following the derivations in the conceptual framework (see Appendix C), we can write

$$\frac{E_{d+1}(T_i)}{E_d(T_i)} = 1 - \frac{var_d(T_i)}{E_d(T_i)\left[1-E_d(T_i)\right]}.$$

Given the mean-preserving spread in the arrival rates, we have

$$E_0(T_i) \approx \bar{\lambda}[1-F(R)],$$

We also have

$$\begin{aligned}
var_0(T_i) &\approx var_0\Big([1-F(R)]\left[d\lambda_i - \kappa d\hat{\lambda}_i\right]\Big), \\
&= [1-F(R)]^2 var_0\left[d\lambda_i - \kappa d\hat{\lambda}_i\right]. \\
&= [1-F(R)]^2\left[var_0(d\lambda) + \kappa^2 var_0\left(d\hat{\lambda}\right) - 2\kappa cov_0\left(d\lambda, d\hat{\lambda}\right)\right]
\end{aligned}$$

where

$$
\begin{aligned}
cov_0\left(d\lambda, d\hat{\lambda}\right) &= B_1\phi\left(\lambda^h\right)^2 + B_1\left(1-\phi\right)\left(\lambda^l\right)^2 + \left(1-B\right)\phi\lambda^h\lambda^l + \left(1-B_1\right)\left(1-\phi\right)\lambda^h\lambda^l - \bar{\lambda}^2 \\
&= B_1 var_0\left(d\lambda\right) + \left(1-B_1\right)\left[\lambda^h\lambda^l - \bar{\lambda}^2\right] \\
&= \left[2B_1 - 1\right] var_0\left(d\lambda\right).
\end{aligned}
$$

The last equality follows since $var_0\left(d\lambda\right) = \bar{\lambda}^2 - \lambda^h\lambda^l$ for $\phi = 1/2$.

Hence,

$$
var_0(T_i) = \left[1 - F(R)\right]^2\left[1 + \kappa^2 - 2\kappa\left[2B_1 - 1\right]\right] var_0\left(d\lambda\right).
$$

Small changes in the dispersion leave the expected job-finding rate unaffected to a first-order, but do increase the variance in job finding rates. However, the increase in the dispersion is scaled and has a smaller impact on the variance in job-finding rates, the higher $B_1$. The first part of the Proposition immediately follows.

Second, we consider the introduction of geometric depreciation of the true and perceived arrival rates,

$$
\begin{aligned}
\lambda_{d+1} =& (1-\theta)\lambda_d, \\
\hat{\lambda}_{d+1} =& (1-B_\theta\theta)\hat{\lambda}_d.
\end{aligned}
$$

We can write,

$$
\begin{aligned}
\frac{T_{d+1}}{T_d} =& (1-\theta)\frac{1-F(R_{d+1})}{1-F(R_d)}, \\
\Rightarrow \frac{d\left[\frac{T_{d+1}}{T_d}\right]}{d\theta} =& -\frac{1-F(R_{d+1})}{1-F(R_d)} + (1-\theta)\frac{d\left[\frac{1-F(R_{d+1})}{1-F(R_d)}\right]}{d\theta}.
\end{aligned}
$$

Unpacking the last term, we find

$$
\begin{aligned}
\frac{d\left[\frac{1-F(R_{d+1})}{1-F(R_d)}\right]}{d\theta} =& \frac{f(R_d)[1-F(R_{d+1})]\frac{dR_d}{d\theta} - f(R_{d+1})[1-F(R_d)]\frac{dR_{d+1}}{d\theta}}{[1-F(R_d)]^2}, \\
=& \frac{f(R_d)\frac{1-F(R_{d+1})}{1-F(R_d)}\frac{dR_d}{d\theta} - f(R_{d+1})\frac{dR_{d+1}}{d\theta}}{1-F(R_d)}, \\
=& \frac{f(R_{d+1})\frac{dR_{d+1}}{d\theta}}{1-F(R_d)}\left[\frac{f(R_d)}{f(R_{d+1})}\frac{1-F(R_{d+1})}{1-F(R_d)}\frac{\frac{dR_d}{d\theta}}{\frac{dR_{d+1}}{d\theta}} - 1\right].
\end{aligned}
$$

We now look at the reaction of the respective reservations wage to the depreciation parameter. The reservation wage is characterized by $V(R_d) = U_d$ where,

$$V(R_d) = \frac{1+\delta}{\delta} u(R_d)$$
$$U_d = u(b_u) + \frac{1}{1+\delta} \max_{R_d} \left\{ U_{d+1} + (1 - B_\theta \theta)^d \lambda_0 \int_{R_d} \big[V(w) - U_{d+1}\big] dF(w) \right\},$$

so substituting the former into the latter for $U_d, U_{d+1}$,and $V(w)$ gives,

$$\frac{1+\delta}{\delta} u(R_d) = u(b_u) + \frac{1}{\delta} \max_{R_d} \left\{ u(R_{d+1}) + (1 - B_\theta \theta)^d \lambda_0 \int_{R_d} \Big[u(w) - u(R_{d+1})\Big] dF(w) \right\}.$$

Total differentiation yields,

$$\frac{1+\delta}{\delta} u'(R_d) dR_d = -\frac{1}{\delta} dB_\theta (1 - B_\theta \theta)^{d-1} \lambda_0 \int_{R_d} \Big[u(w) - u(R_{d+1})\Big] dF(w) d\theta \ldots$$
$$\ldots + \frac{1}{\delta} u'(R_{d+1}) \frac{dR_{d+1}}{d\theta} d\theta - \frac{1}{\delta} (1 - B_\theta \theta)^t \lambda_0 u'(R_{d+1}) \frac{dR_{d+1}}{d\theta} d\theta,$$

Hence, we find

$$\frac{dR_d}{d\theta} = \frac{1}{1+\delta} \left\{ -d \frac{B_\theta}{1 - B_\theta \theta} \left( \frac{1 - B_\theta \theta}{1 - \theta} \right)^d T_d E\left[ \frac{u(w) - u(R_{d+1})}{u'(R_d)} \middle| w > R_d \right] + \frac{u'(R_{d+1})}{u'(R_d)} (1 - \hat{\lambda}_d) \frac{dR_{d+1}}{d\theta} \right\},$$

and, then by iterating, we get

$$\frac{dR_d}{d\theta} = -\frac{1}{1+\delta} \frac{B_\theta}{1 - B_\theta \theta} \sum_{s=d}^{\infty} \left\{ \left( \frac{\prod_{k=d}^{s} [1 - \hat{\lambda}_k]}{1 - \hat{\lambda}_s} \right) \frac{u'(R_{s+1})}{u'(R_d)} s \left( \frac{1 - B_\theta \theta}{1 - \theta} \right)^s T_s E\left[ \frac{u(w) - u(R_{s+1})}{u'(R_s)} \middle| w > R_s \right] \right\}.$$

Starting from $\theta \approx 0$, the reservation wage, arrival rate, and job-finding rate are approximate constant and the perceived arrival rate equals the actual arrival rate. Denoting by $R$ and $T = \lambda [1 - F(R)]$ the reservation wage and the job finding for the stationary type, where we have dropped the subindex 0 in the notation, we can write

$$\left. \frac{dR_{d+1}}{d\theta} \right|_{\theta=0} = -\frac{1}{1+\delta} B_\theta T E\left[ \frac{u(w) - u(R)}{u'(R)} \middle| w > R \right] \sum_{s=d+1}^{\infty} \left\{ (1-\lambda)^{s-d-1} s \right\},$$

and thus

$$\left. \frac{\frac{dR_d}{d\theta}}{\frac{dR_{d+1}}{d\theta}} \right|_{\theta=0} = \frac{\sum_{s=d-}^{\infty} (1-\lambda)^{s-d} s}{\sum_{s=d+1}^{\infty} (1-\lambda)^{s-d-1} s} = \frac{d + (1-\lambda) \sum_{s=d+1}^{\infty} (1-\lambda)^{s-d-1} s}{\sum_{s=d+1}^{\infty} (1-\lambda)^{s-d-1} s},$$
$$= \frac{d + (1-\lambda) \left[ \frac{d+1}{\lambda} + \frac{1-\lambda}{\lambda^2} \right]}{\frac{d+1}{\lambda} + \frac{1-\lambda}{\lambda^2}} < 1,$$

which proves that the reservation wage responds more at longer durations. The last equality above follows from expanding the power series as follows:

$$\begin{aligned}
\sum_{s=d+1}^{\infty}(1-\lambda)^{s-d-1}s =& d+1+(1-\lambda)(d+2)+(1-\lambda)^2(d+3)+(1-\lambda)^3(d+4)+\ldots,\\
=&(d+1)(1+(1-\lambda)+(1-\lambda)^2+(1-\lambda)^3+\ldots)+(1-\lambda)+2(1-\lambda)^2+\ldots,\\
=&\frac{d+1}{\lambda}+(1-\lambda)(1+(1-\lambda)+(1-\lambda)^2+(1-\lambda)^3+\ldots)+(1-\lambda)^2+2(1-\lambda)^3+\ldots,\\
=&\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda}+(1-\lambda)^2(1+(1-\lambda)+(1-\lambda)^3+\ldots)+(1-\lambda)^3+2(1-\lambda)^4+\ldots,\\
=&\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda}+\frac{(1-\lambda)^2}{\lambda}+\frac{(1-\lambda)^3}{\lambda}+\frac{(1-\lambda)^4}{\lambda}+\ldots,\\
=&\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda}(1+(1-\lambda)+(1-\lambda)^2+(1-\lambda)^3+\ldots),\\
=&\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda^2}.
\end{aligned}$$

Hence, putting things together and starting from $\theta \approx 0$, we have

$$\begin{aligned}
\left.\frac{d\left[\frac{T_{d+1}}{T_d}\right]}{d\theta}\right|_{\theta=0} =& -1+\frac{f(R)\left.\frac{dR_{d+1}}{d\theta}\right|_{\theta=0}}{1-F(R)}\left[\left.\frac{\frac{dR_d}{d\theta}}{\frac{dR_{d+1}}{d\theta}}\right|_{\theta=0}-1\right],\\
=&-1+\frac{f(R)}{1-F(R)}\frac{1}{1+\delta}B_\theta TE\left[\frac{u(w)-u(R)}{u'(R)}\middle| w>R\right]\left\{\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda^2}\right\}\ldots\\
&\ldots\left[1-\frac{d+(1-\lambda)\left[\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda^2}\right]}{\frac{d+1}{\lambda}+\frac{1-\lambda}{\lambda^2}}\right],\\
=&-1+\frac{f(R)}{1-F(R)}\left[1+\frac{1-\lambda}{\lambda}\right]\frac{1}{1+\delta}B_\theta TE\left[\frac{u(w)-u(R)}{u'(R)}\middle| w>R\right],\\
=&\frac{1}{1+\delta}B_\theta E\left[\frac{u(w)-u(R)}{u'(R)}\middle| w>R\right]f(R)-1,\\
=&B_\theta\times\frac{\kappa}{\lambda}-1.
\end{aligned}$$

Moreover, since $\frac{dR}{dB_\theta}=0$ for $\theta=0$, we also have

$$\left.\frac{d^2\left[\frac{T_{d+1}}{T_d}\right]}{d\theta dB_\theta}\right|_{\theta=0}=\frac{\kappa}{\lambda}>0.$$

This proves the second part of the Proposition.

## F.2 Numerical Analysis

Table F1 shows the 8 moments that we target in the calibration of our structural model. As in the statistical model, the targeted moments include the actual and perceived job-finding rates for the short, medium and long-term unemployed. We additionally target an average job acceptance rate underlying the job-finding rates of 0.71, as estimated by Hall and Mueller [2018] using the KM survey. As we already estimated the true duration dependence in our statistical model using elicited beliefs moments, instead of targeting these again, we directly target a moment capturing the true depreciation in job finding, i.e., the average of the ratio of true job finding when long-term vs. short-term unemployed within a spell (i.e., $E_{7+}(T_{i,d})/E_{06}(T_{i,d})$ for a given spell). We simulate this moment using the baseline estimation of our statistical model, obtaining a value of 0.895. We also gauge the robustness of our results to the rate of depreciation and recalibrate the model targeting a ratio 0.75, which is below any estimate we obtain in the statistical model (excluding the specification in which we do not allow for heterogeneity). We set the perceived true duration dependence $B_\theta$ equal to 0 in both specifications.

Table F2 Panel A shows the set parameter values. We set the separation rate at 0.02 per month, corresponding the average separation rate in the SCE. We set arrival rate of job offers for employed workers at 0.15, in line with recent evidence in Faberman et al. [2017] also using the SCE. We assume that wages are log-normally distributed, with a standard deviation of the logged distribution of $\sigma_w = 0.24$ as estimated by Hall and Mueller [2018] with the KM survey data. We normalize the median of the wage offer distribution to 1. We also assume an annual discount factor 0.996 and CRRA preferences with relative risk aversion equal to 2. Panel B of Table F2 shows the remaining 7 parameters of our model $\{B_0, B_1, \lambda_l, \lambda_h, \phi, \theta, b_u\}$ that are estimated by targeting the vector of 8 moments. We assume that the flow utility when unemployed $u(b_u)$ remains constant throughout the spell. We also assume that the arrival rates depreciate at geometric rate $\theta$ for the first 24 months of the spell, but then remain constant so that the unemployment state becomes stationary. We assume that a worker who is separated after an employment spell of less than six months starts the unemployment spell in this stationary state.

The estimated parameters minimize the sum of squared differences between data moments and simulated moments from the model. We find that the uniform bias parameter $B_0$ is negative, but the average bias is still positive. This is due to the share of low types perceiving themselves as high, who remain unemployed for the longest. The probability that high (low) types perceive themselves as high (low) types equals $B_1 = 0.84$ in the baseline specification. As we assume that none of the true duration dependence in job finding is perceived ($B_\theta = 0$), the corresponding cross-sectional bias becomes smaller in the model where we target high true duration dependence ($B_1 = 0.89$).[3]

Table F1 shows that we closely match our targeted moments. We also obtain plausible values for standard labor market statistics; the elasticity of the unemployment duration with respect to unemployment benefits is 0.62, which is within the range of estimates in the literature (see Schmieder and von Wachter [2016]). The elasticity of the reservation wage equals 0.59, which corresponds to the estimate in Fishe [1982]. The elasticity of the equilibrium wage equals .04, which is arguably low, but still higher than recent estimates in Jaeger et al. [2020]. The monthly rate of job-to-job transitions equals 0.024,

[3] We have also extended our model with a type-specific bias in the perceived arrival rates. This relaxes the restrictions of our stylized model that on average the low-type job seekers are more optimistic than the high-type job seekers. However, the estimated type-specific biases are very close, suggesting that this restriction is not binding.

which is within the range considered by Hornstein et al. [2011].[4]

Table F3 shows the impact of eliminating the biases in beliefs on the average unemployment durations and the share of long-term unemployed. The intermediate columns consider the elimination of one bias at a time, the last column the elimination of all biases simultaneously. From Panel A, which shows the results for the baseline model, we see that eliminating all biases lowers the average unemployment duration, but this effect is numerically very small. Despite the small impact on the overall duration, the impact on the share of LT unemployed is substantial, which decreases by 8.1 percent (2.4 percentage points) when all biases are eliminated. Panel B shows the results for the model calibrated with high depreciation rate. The effect on the average unemployment duration is somewhat larger, at around 0.15 months. However, eliminating the biases reduces the share of LT unemployed by 2.6 percentage points, which is slightly higher than in the baseline model. Overall, the model's prediction that biased beliefs contribute substantially to the high incidence of LT unemployment is robust to the relative importance of heterogeneity vs. true depreciation in the arrival rates.[5]

[4]We also performed sensitivity checks when changing incidental parameters, including the dispersion of the wage distribution, the level of risk aversion, the arrival rate of job offers for the employed workers and for the separated workers after short employment spells, which all change the relative value of unemployment to employment. For the baseline model, it is mainly the parameter $b_u$ affecting the flow value of unemployment that adjusts, while the other parameter estimates remain very similar. The other parameter estimates become more sensitive in the model with high depreciation.

[5]We note that these counterfactual results remain very similar when changing incidental parameters (i.e., wage offer distribution, arrival rates, risk aversion) in the baseline calibration, but are somewhat sensitive in the calibration with high depreciation.

Table F1: Targeted Data Moments and Corresponding Moments in Structural Model

| Moments | Data | Baseline Model | High-Depreciation Model |
|---|---|---|---|
| Average of Realized 3-Month Job-Finding Rates: | | | |
| ... at 0-3 Months of Unemployment | 0.642 | 0.641 | 0.640 |
| ... at 4-6 Months of Unemployment | 0.472 | 0.470 | 0.474 |
| ... at 7 Months of Unemployment or more | 0.256 | 0.259 | 0.255 |
| Average of Elicited 3-Month Job-Finding Probability: | | | |
| ... at 0-3 Months of Unemployment | 0.616 | 0.614 | 0.614 |
| ... at 4-6 Months of Unemployment | 0.529 | 0.537 | 0.534 |
| ... at 7 Months of Unemployment or more | 0.409 | 0.404 | 0.405 |
| Acceptance Rate: | 0.710 | 0.715 | 0.715 |
| True Duration Dependence: | | | |
| ... Baseline Depreciation | 0.895 | 0.895 | - |
| ... High Depreciation | 0.75 | - | 0.750 |

*Notes*: Survey weights are used for all data moments from the SCE (averages of realized and elicited job finding. The SCE sample is restricted to unemployed workers, ages 20-65, and includes only interviews that were followed by three consecutive monthly interviews. The target for the acceptance rate is from Hall and Mueller [2018], and the target for the true duration dependence in the baseline model is based on the estimates in the statistical model. More precisely, we target the ratio of the sample average of job finding when long-term unemployed ($> 6$ months) vs. the sample average of job finding when short-term unemployed ($\leq 6$ months). This ratio is estimated to be 0.90 in the statistical model. We gauge the sensitivity of our results to setting this target ratio at 0.75 in the high-depreciation model, which is below any estimate from the statistical model.

Table F2: Calibrated Parameters

| Parameters | Symbol | Baseline Model | High-Depreciation Model |
|---|---|---|---|
| **Panel A. Set Parameters** | | | |
| Median of wage offer distribution | $\mu_w$ | 1 | 1 |
| Std. dev. of logged wage offer distribution | $\sigma_w$ | 0.24 | 0.24 |
| Exogeneous job loss probability | $\sigma$ | 0.02 | 0.02 |
| Arrival rate when employed | $\lambda^e$ | 0.15 | 0.15 |
| Discount rate | $\delta$ | 0.004 | 0.004 |
| Coefficient of relative risk aversion | $\gamma$ | 2 | 2 |
| Longitudinal bias | $B_\theta$ | 0 | 0 |
| **Panel B. Estimated Parameters** | | | |
| Uniform bias | $B_0$ | -0.014 | -0.051 |
| Cross-sectional bias | $B_1$ | 0.84 | 0.89 |
| Low-type arrival rate | $\lambda_l$ | 0.12 | 0.16 |
| High-type arrival rate | $\lambda_h$ | 0.63 | 0.67 |
| Share of high-types | $\phi$ | 0.78 | 0.74 |
| True depreciation in arrival rate | $\theta$ | 0.019 | 0.046 |
| Unemployed consumption | $b_u$ | 0.52 | 0.52 |

*Notes*: The table shows the calibrated parameter values, where the parameters in Panel A are normalizations or set based on external information and Panel B shows the estimated parameters that are obtained from targeting the data moments in Table F1.

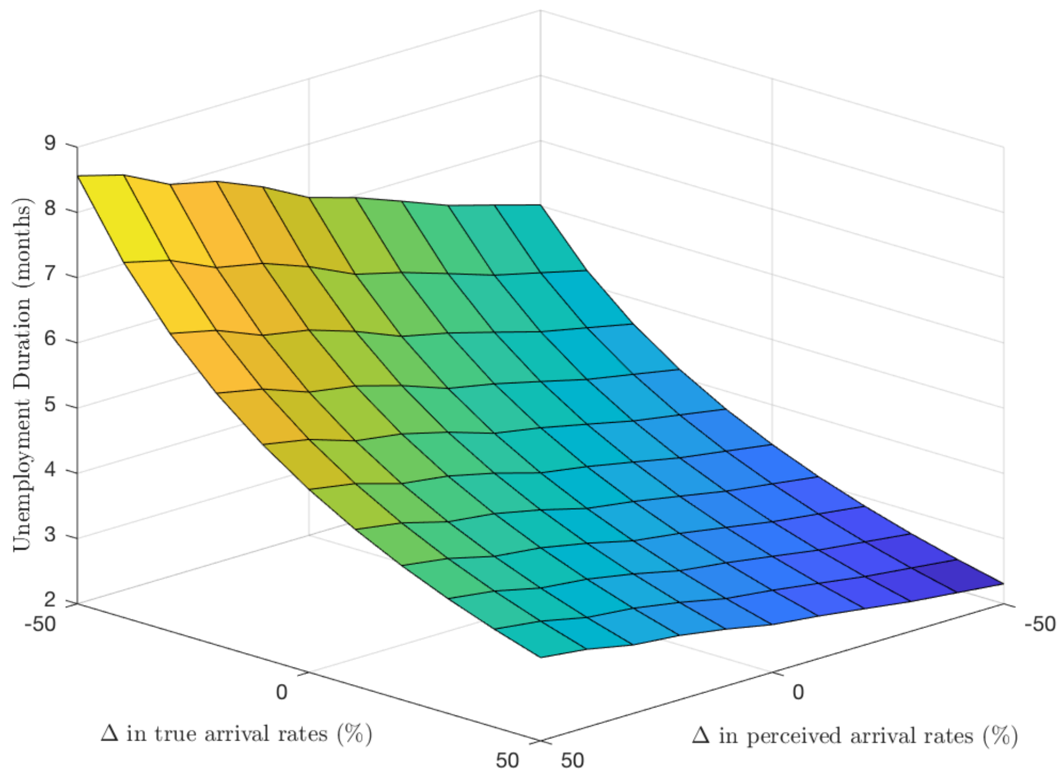
Table F3: Comparative Statics in Structural Model

| | Calibrated Model | Eliminating Biases | | | |
|---|---|---|---|---|---|
| | | $B_0 = 0$ | $B_1 = 1$ | $B_\theta = 1$ | $B_0 = 0$ $B_1 = 1$ $B_\theta = 1$ |
| **Panel A. Baseline Model** | | | | | |
| Unemployment duration | 4.01 | 4.07 | 4.01 | 3.90 | 3.95 |
| Share of LT unemployed | 0.295 | 0.297 | 0.275 | 0.289 | 0.271 |
| **Panel B. High-Depreciation Model** | | | | | |
| Average unemployment duration | 4.00 | 4.15 | 3.98 | 3.74 | 3.86 |
| Share of LT unemployed | 0.293 | 0.299 | 0.278 | 0.278 | 0.267 |

*Notes*: The table reports selected moments for the baseline calibration of the structural model (Panel A), a calibration with a higher individual-level depreciation in true job-finding probabilities (Panel B) and counterfactual simulations where the biases in the respective model are eliminated. Besides the depreciation in job finding, both calibrations match the same set of moments (see Table F1), resulting in the same average unemployment duration and share of long-term unemployed.

Figure F1: Comparative Statics: True vs. Perceived Changes in Arrival Rates

**A. Impact of Arrival Rates on Duration**




**B. Impact of Heterogeneity on LT Incidence**

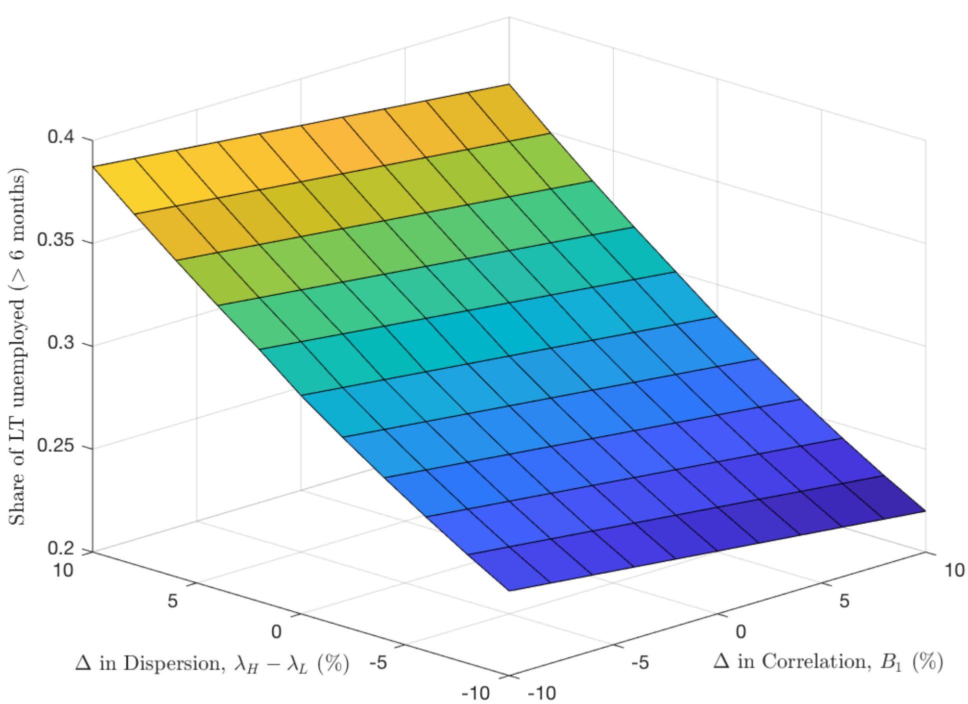
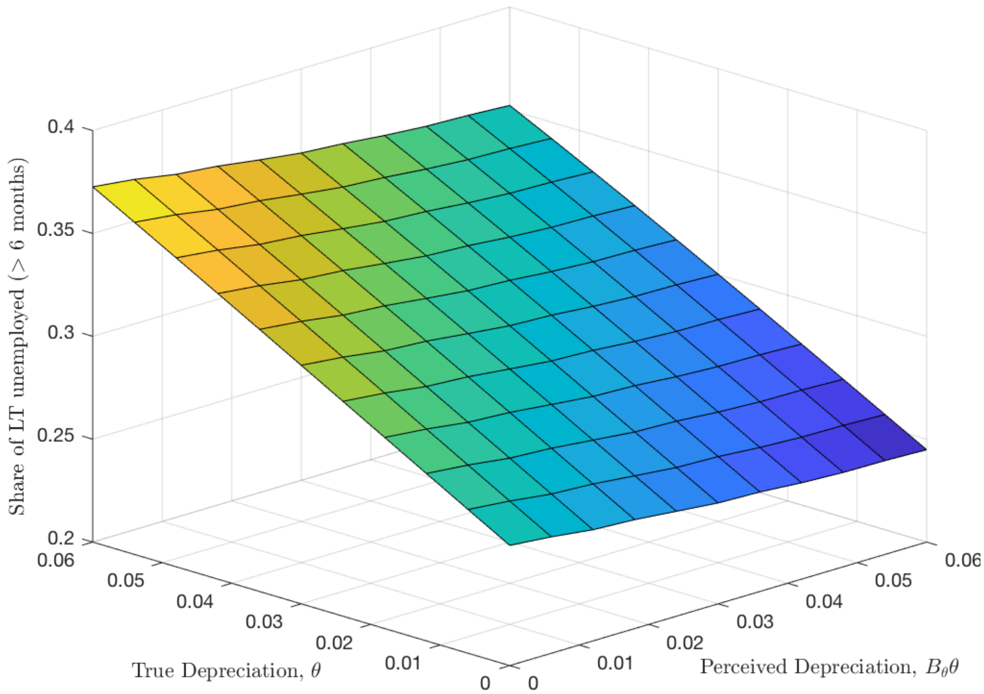


Figure F1: Comparative Statics: True vs. Perceived Changes in Arrival Rates (*continued)*

**C. Impact of Depreciation on LT Incidence**




*Notes:* Panel A plots the average unemployment duration as a function of actual and perceived arrival rates, changing them in the same way for all types relative to the baseline model. Panel B plots the share of long-term unemployment (i.e., the share of unemployed workers who are unemployed for longer than 6 months) as a function of the spread of true arrival rates (while preserving the mean arrival rate) and the correlation between the perceived and true arrival rates. Panel C plots the share of long-term unemployment as a function of the true and perceived depreciation rate. The output in the last two panels corresponds to results (i) and (ii) in Proposition 1.