Factors Determining Callbacks to Job Applications: An Audit Study

Henry S. Farber – Princeton University
Daniel Silverman – Arizona State University
Till von Wachter – UCLA

SOLE Presidential Address
### May 6, 2016

Outline of Talk

- **Introduction**
- Bayesian Updating Model
- Design of Audit Study
- Univariate Results
- Multivariate Results
- Why are findings different from Kroft-Lange-Notowidigdo?
- Conclusion

Mean and Median Duration of Unemployment Spells in Progress
# CPS Data by Quarter

Source: Current Population Survey
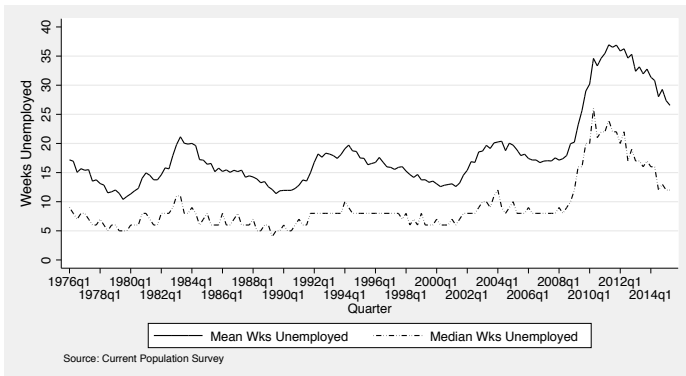
- Long-term unemployment remains a problem in aftermath of Great Recession.
- Mean and median duration at end of 2014 are above levels seen prior to Great Recession.

Job Finding Probability by Unemployment Duration, 2008-2014

Source: Matched Monthly Current Population Survey
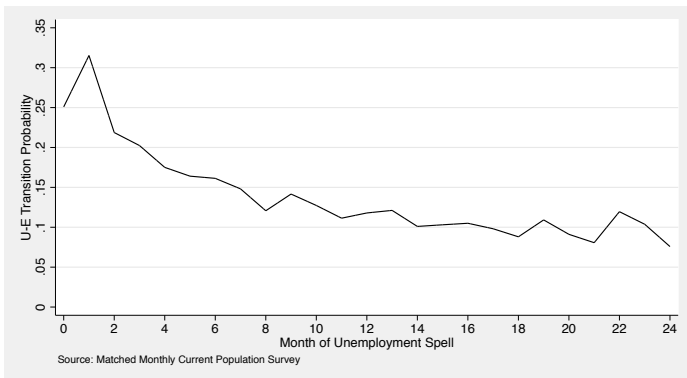
- Job finding rates decline with unemployment duration
- Among others due to skill depreciation or signalling
- Yet, relationship is not necessarily causal because of selection in who exits across the unemployment spell.

Mean Duration of Unemployment Spells in Progress, by Age

Average Duration of Unemployment, by Age

Source: Current Population Survey
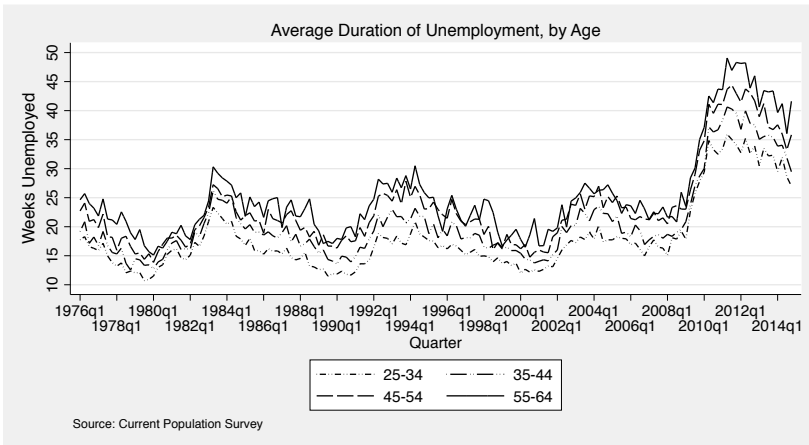
- Unemployment durations increase sharply with age.
- Particularly large difference by age since Great Recession.
- Highlights difficulties of older job losers.

Existing Literature on Effect of Unemployment Duration on Callback

- Few studies of causal effect of long-term unemployment (LTU) on exit hazard, wages, or other outcomes
  - ⇒ Difficult to address selection problem without exogenous variation in durations, which is rarely available
- High-profile audit study by Kroft, Lange, and Notowidigdo (2013) finds negative effect of unemployment duration on resume on call-back rates for younger workers.
  - KLN also find that effect smaller in high unemployment cities, suggesting that negative effect of LTU is due to a signal (rather than a negative effect on productivity)
- Other audit study evidence is mixed.
  - Ghayad (2014) –negative effect of U duration in first 6 months.
  - Nunley, et al (2014) – no effect of unemployment duration.
  - Eriksson and Rooth (2014) – no effects in Sweden for college grads and no effect before 6 months for less skilled.

Audit Studies, Effect of Unemployment History on Callback

Study	Locus	Age & Sex	Educ	Occupation	Months Unemp. Dur (UD)	UD Effect on Callbacks
Oberholzer-Gee (2008)	Zurich, 1999	25-29 F	–	Admin Asst	6-30	$+ UD < 12$, $- UD > 18$.
Kroft et al. (2013)	US 100 MSAs, 2011-12	19-40 M&F	HS, AS, BA	Admin, Cler, Cust. Serv., Sales	1-36	- effect tight LM. Weaker - in slack LM.
Ghayad (2013)	US, 25 MSAs, 2012	25-30 M	BA	Prof, Sales, Admin. Supp.	1-12	- effect
Eriksson and Rooth (2014)	Sweden 2 cities	20-30 M&F	Various Skill	Various (13)	0-9	- effect low skill, $UD \geq 9$. 0 other.
Nunley et al. (2014)	2013, US	25-26 M&F	BA	Marketing, Sales, Assistant Manager	3-12	0 effect
Farber et al. (2015)	US, 8 MSAs, 2012-14	35-58 F	BA	Admin, Cler	1-12	0 effect

Existing Literature on Effect of Age on Callback

- General evidence that job loss is very costly for older workers.
- Unknown if long-term unemployment is particularly costly for older workers.
- Only a small literature on the effect of age on the callback rate.
  - Lahey (2008) finds large negative effects of age on the callback rate for women seeking entry-level positions in the U.S.
  - Neumark, et al (2015) (a new audit study) finds strong age effects for women but not so much for men.

Activity During Unemployment Spell – Interim Job

- Individuals may need to take an "interim job" during their unemployment spell in order to meet liquidity needs.
- They may be jobs for which the individual appears ill-matched in the sense of being lower-skilled than the jobs the individual has held in the past.
  - Such an interim job for an individual seeking permanent employment in mid-level office administration might be as a restaurant server
- Potentially opposing effects of holding an interim job on likelihood of being hired in an "appropriate" job.
  - (Positive) Potential employer might infer that an applicant holding an interim job that the individual is energetic, ambitious, and hard-working.
  - (Negative) Potential employer might infer that an applicant holding an interim job is not an appropriate match (myopia in ignoring earlier experience).

Outline of Talk

- Introduction
- **Bayesian Updating Model**
  - **Skip Formal Presentation**
  - **But Will Mention as Appropriate**
- Design of Audit Study
- Univariate Results
- Multivariate Results
- Why are findings different from Kroft-Lange-Notowidigdo?
- Conclusion

Information, Updating, and the Callback Decision

- Assume a profit-maximizing risk-neutral firm with a single employee whose output (Y) is equal worker quality (μ).
- Assume all potential workers will be paid the same wage.
  - $=>$ Firm wants to hire most able worker among applicants.
- The firm has a noisy signal of μ derived from an application.
- This noisy signal has many components (e.g., work history, education, unemployment experience, age).
- Think of the employer as using Bayesian updating to derive an expectation about applicant quality based on all available information.

Learning About Worker Quality: Notation

- The firm has incomplete information about the quality (μ_i) of applicant i.
- Makes inference about μ_i based on a set of k noisy signals.
  - For our purposes, signals include among other background information, the applicant's unemployment experience, age, and whether the applicant holds an interim job.
- s_{ij} represent the jth noisy signal of μ_i:

$$s_{ij} = \frac{1}{\alpha_j}\mu_i + \gamma_{ij}.$$

- $\gamma_{ij} \sim N(0, \sigma_j^2)$.
- Parameters α_j are normalizations that account for fact that some signals are positive and some negative as well as for differential scaling of the signals.

Employer Updating Based on Signals

- The employer's inference problem is to combine the available information on $s_{ij}, j = 1, ..., k$ optimally in order to derive an expected value for applicant quality ($E(\mu_i|s_{i1}, ..., s_{ik})$).
- The precision of each signal is the inverse of its variance: $h_j = 1/\sigma_j^2$.
- Bayesian posterior beliefs are straightforward precision weighted average of the signals.
- The posterior expectation is

$$E(\mu_i|s_{i1}, ..., s_{ik}) = \frac{\sum_{j=1}^{k} h_j \alpha_j s_{ij}}{\sum_{j=1}^{k} h_j}.$$

Marginal Effect of Signal on Expected Worker Quality

- Marginal effect of a change in signal m (s_{im}) on expected worker quality is

$$\frac{\partial E(\mu_i)}{\partial s_{im}} = \alpha_m \left[\frac{h_m}{\sum_{j=1}^{k} h_j} \right]$$

- Takes the sign of α_m.
    - If signal m is unemployment duration then $\alpha_m < 0$, and the marginal effect of unemployment duration is negative.
    - If signal m is age and age is a negative signal of worker quality, then $\alpha_m < 0$ and older workers have lower posterior mean worker quality.
    - If signal m is the holding of a low-level interim job, there is no clear prediction regarding the sign of α_m.

Effect of U Duration Varies with State of Labor Market

- Likely there is more information about applicant quality in unemployment duration when labor market is tighter.
- Formally, the precision associated with the unemployment duration signal is higher where the local unemployment rate is lower so that there is relatively more updating based on unemployment duration.

$$\frac{\partial^2 E(\mu_i)}{\partial s_{im} \partial h_m} = \alpha_m \left[\frac{1}{\sum_{j=1}^k h_j} \right] \left[1 - \frac{h_m}{\sum_{j=1}^k h_j} \right]$$

- This expression has the sign of α_m.
- $\alpha_m < 0$ where s_m represents unemployment duration.
- => The negative marginal effect of unemployment duration on the likelihood of callback is larger in absolute value in tighter labor markets.
    - Found by Kroft, Lange, and Notowidigdo.

More Signals => Any One Signal Less Important

- Older workers / longer experience => more information
  => increase the number of signals (k).

$$\frac{\partial E(\mu_i)}{\partial s_{im}} = \alpha_m \left[\frac{h_m}{\sum_{j=1}^{k} h_j} \right]$$

- An increase in k increases the denominator.
- => A reduction in the absolute value of the marginal effect any particular existing signal.
- For example, the marginal effect of unemployment duration will be smaller for older workers.
  - Intuitively, older workers have a longer employment history that will dilute the effect of recent unemployment on the likelihood of callback.
- Generally, callback rates for older workers should be less affected by particular elements.

A Final Prediction: Employer Selectivity

- Not based strictly on updating model.
- If an employer has a great need for workers then the employer may not be as selective.
  - Indicated by a higher callback rate for applicants.
  - Threshold posterior mean worker quality necessary for a callback will be lower where demand is high.
- Implication is that marginal effect of particular worker attributes (unemployment duration, age, and the holding of a low-level interim job in this case) on the likelihood of callback will be smaller in absolute terms for less selective employers.

The Role of Specific Components of the Information Set

- How much any component affects the inference about μ depends on two things:
  1. The precision with which the component provides information about μ.
  2. The precision with which other components provide information about μ.

- The usual Bayesian posterior mean of μ is a weighted average of the prior means of μ conditional on specific components where the weights are proportional to the variances of the priors on μ based on each component.

The Role of Unemployment Duration in the Updating Process

- Employers may infer from a long U spell that the worker is not good $=>$ lower posterior mean μ and lower probability of callback.
- If labor market weak, there may be less information in the long U signal (high variance) $=>$ smaller effect on the posterior mean of μ and smaller effect on callback probability.
  - This is a result found by KLN, and it motivated our choice of 4 low U and 4 high U cities.
  - We do not follow this up given our baseline results (below).
- If labor market has more information about the worker in other dimensions, $=>$ smaller effect of the long U signal on the posterior mean of μ and smaller effect on callback probability.

The Role of Age in the Updating Process

- Employers may value older workers more highly if they have more human capital or other attributes that employers value
    - $=>$ higher posterior mean $\mu =>$ higher callback probability.
- Employers may value older workers less if they have less relevant human capital, short time horizons or other attributes that employers do not like.
    - $=>$ lower posterior mean μ and lower probability of callback.
- The resume of an older worker has more information than that of a younger worker due to longer work history.
    - $=>$ Other particular factors (e..g., duration of U or interim job) may have smaller effects on the posterior mean of μ for older workers because the prior mean is relatively low variance.
- Bottom Line: No clear prediction on age.

The Role of a Low-Skill Interim Job in the Updating Process

- Employer may view a worker with even a low-skill interim job as a "go-getter."
    - $=>$ higher posterior mean μ and higher probability of callback.
- Employer may view a worker with a low-skill interim job as an inappropriate match despite earlier experience.
    - $=>$ lower posterior mean μ and lower probability of callback.
    - It may be that the kind of mechanized first-stage screening of resumes that has become common automatically put a lower score on resumes with inappropriate worker experience.
- Bottom Line: No clear prediction on interim job.

Outline of Talk

- Introduction
- Bayesian Updating Model
- **Design of Audit Study**
- Univariate Results
- Multivariate Results
- Why are findings different from Kroft-Lange-Notowidigdo?
- Conclusion

Overview of Approach and Findings

Snapshot of Our Approach:

- Implemented an audit study varying resume characteristics for mature & older women applying for clerical jobs
- Varied three aspects of resume
  1. Duration of unemployment spell
  2. Age implied by resume
  3. Presence of low-quality interim jobs

Four main findings on callback rates:

1. Find no effect of unemployment duration on callback.
2. Confirm negative effect of age on callback.
3. Find negative effect of low-quality interim jobs on callback.
4. Find that firms with higher call back rates are less sensitive to age, interim jobs.

Audit Study of Likelihood of Callback to Job Applications

- Uses technique pioneered by Bertrand and Mullainathan (2004) to study effect of typically African-American names on response to job applications.
- Basic idea is to submitted carefully constructed job applications to on-line job postings, vary key aspects of the applications, and measure variation in callback rates.
- We vary three features of applications:
  1. Duration of unemployment
  2. Age
  3. Employment in a low-skill interim job while searching for a higher skill job that matches previous work experience.
- In our case, to better target resumes to job postings, we limited our applications to administrative/clerical jobs.

Design of Experiment

- To further limit heterogeneity, All applicants shared a set of attributes:
  - Female
  - Four-year college graduates
  - Experience in low- to mid-level administrative/office work
- Selected 8 U.S. cities – 4 low and 4 high U rate

Low U	2012	2014	High U	2012	2014
Dallas, TX	6.6	5.0	Charlotte, NC	9.2	6.0
Omaha, NE	4.4	3.7	Chicago, IL	9.1	7.0
Pittsburgh, PA	7.2	5.6	Sacramento, CA	10.3	7.2
Portland, ME	6.1	4.6	Tampa, FL	8.3	6.1
Average	6.1	4.7	Average	9.2	6.6

Design of Experiment, continued

- Varied specific characteristics of applications:
  - Unemployment duration 0 to 52 weeks.
  - Age 35-58
  - Interim job at end of unemployment spell – Yes or No.
    - Typical Interim jobs were sales associate or cashier at a big box or grocery store, and restaurant server.
- Experiment carried out in 4 rounds

Round	Dates
1	March-May 2012
2	July-September 2012
3	November 2013 - April 2014
4	April 2014 - August 2014

- Precise structure of randomization varied by round as our experience evolved. – Described below.

Design of 4 Rounds: Summary I

- Number of applications per posting:
  - Rounds 1-3 have 2 applications/posting.
  - Round 4 has 4 applications per posting.
- Variation in unemployment duration:
  - Within-posting variation in U duration in all rounds.
  - Round 1
    - "Control" applicant to each job had just entered unemployment (0 weeks unemployed).
    - "Treatment" applicant randomly drawn from set of 4, 12, 24, 52 weeks unemployed.
  - Rounds 2-4
    - Randomly assigned unemployment duration to *both* applicants for each job posting.
    - Unemployment durations drawn from set of 0, 4, 12, 24, 52 weeks unemployed.

Design of 4 Rounds: Summary II

- **Variation in age:**
  - Rounds 1-3 have no within-posting variation in age but random variation across postings.
  - Round 4 has 2 "young' and 2 "old" applications per posting, randomly assigned.
- **Variation in Interim jobs:**
  - Rounds 1 and 2 have no interim jobs.
  - Rounds 3 and 4 have independent random assignment of and interim jobs within posting.

Basic Statistics of Audit Study – Analysis Sample

Number of Job Openings Per Round:
- Round 1 (March-May 2012): 1,027 job openings
- Round 2 (July-September 2012): 1,215 job openings
- Round 3 (November 2013-April 2014): 834 job openings
- Round 4 (April-August 2014): 1,518 job openings

Basic Statistics Promising:
- Analyze callback rates from 12224 resumes sent to 4594 job postings.
- Reasonable mean call back rate of 10.37 percent.
- Higher mean callback rate in cities with low unemployment rates: (12.2 percent vs. 8.9 percent).

Next Assess Effect of Each of 3 Randomized Resume Characteristics Separately on Callback Rates

Outline of Talk

- Introduction
- Bayesian Updating Model
- Design of Audit Study
- **Univariate Results**
  - **Not Presented Today**
  - **Same Findings as Multivariate Analysis**
- Multivariate Results
- Why are findings different from Kroft-Lange-Notowidigdo?
- Conclusion

Summary of Univariate Results

Four main findings on callback rates:

1. Find no effect of unemployment duration on callback.
2. Find negative effect of age on callback.
   - About 2-3 percentage points lower for 55 and older relative to 35-54.
3. Find negative effect of low-quality interim job on callback.
   - About 1.5 percentage points lower for interim job holders.
4. Find that firms with higher call back rates are less sensitive to age, interim jobs.

Outline of Talk

- Introduction
- Bayesian Updating Model
- Design of Audit Study
- Univariate Results
- **Multivariate Results**
- Why are findings different from Kroft-Lange-Notowidigdo?
- Conclusion

Multivariate Analysis Pooling 4 Rounds & All Treatments:

1. Begin by Using Logit Model of Probability of a Callback.
   - Results hold independent of specific functional form.
   - Want to have consistent model for incorporating job effects.
2. Account for job opening diffs w/Random-Effects Logit Model.
   - Random-effects assumption relevant for our case, because job applications were sent out randomly.
   - Improve precision by accounting for random variation in callback rates across jobs.
   - Presence of random effect can affect point estimates through non-linearity.
3. Account for job-opening diffs using Fixed Effects Logit (Chamberlain's conditional logit).
   - Conditions on number of successes.
   - Appropriate if we are worried about correlation of the fixed effects with the treatment (Unlikely in our case by design).
   - Approach automatically drops those jobs in which there is no within-opening variation in callback rate.

Logit, Random Effects Logit, and Conditional Logit Estimates: Odds Ratios

Variable	Logit	RE Logit	FE Logit
U Duration			
4 Weeks U	0.973	0.948	0.951
	(0.084)	(0.139)	(0.151)
12 Weeks U	1.140	1.260	1.278
	(0.100)	(0.181)	(0.203)
24 Weeks U	1.084	1.158	1.170
	(0.092)	(0.164)	(0.182)
52 Weeks U	0.990	1.111	1.178
	(0.086)	(0.163)	(0.188)
Age 55-58	0.791	0.566	0.531
	(0.055)	(0.059)	(0.063)
Interim Job	0.850	0.725	0.715
	(0.065)	(0.088)	(0.092)
N Apps.	12224	12224	1604

Note: U duration = 0 is base category. Robust standard errors clustered by job id in parentheses. Logit and RE logit estimates also control for round and whether the city was low-unemployment.

- Hausman test does not reject H_0 that RE logit consistent (p-value=0.84).

- Results basically the same as univariate analyses.

Confirm Findings:

- No effect of U duration on callback.

- Older workers less likely to be called back.

  - Odds 43% lower
    => Big effect.

- Interim job reduces likelihood of callback.

  - Odds 27% lower
    => Big effect.

Logit Estimates for Round 4 by Number of Callbacks: Odds Ratios

Variable	All	1-3 Cbacks	1 Cback	2 Cbacks	3 Cbacks
4 Week U	1.092	1.119	0.761	1.470	1.400
	(0.152)	(0.225)	(0.248)	(0.524)	(0.850)
12 Weeks U	1.206	1.450	1.113	1.702	1.692
	(0.181)	(0.306)	(0.326)	(0.732)	(1.058)
24 Weeks U	1.243	1.287	1.227	1.525	1.564
	(0.174)	(0.253)	(0.363)	(0.561)	(1.004)
52 Weeks U	1.092	1.314	0.925	2.081	1.179
	(0.154)	(0.261)	(0.283)	(0.783)	(0.698)
Age 55-58	0.687	0.481	0.373	0.363	0.965
	(0.056)	(0.076)	(0.086)	(0.132)	(0.447)
Interim Job	0.839	0.758	0.476	1.016	1.423
	(0.073)	(0.096)	(0.098)	(0.217)	(0.575)
Log L	-1840.8	-712.1	-326.9	-223.0	-84.5
P-Value (U=0)	0.58	0.43	0.67	0.43	0.93
Sample Size	6072	1092	600	340	152

Note: Numbers in parentheses are standard errors clustered by job id.

- Confirms earlier results.

- Negative effect of age, interim job.

- Effects smaller with higher cback rate.

- No effect of U duration.

Outline of Talk

- Introduction
- Bayesian Updating Model
- Design of Audit Study
- Univariate Results
- Multivariate Results
- **Why are findings different from Kroft-Lange-Notowidigdo?**
- Conclusion

Comparison of Results to Kroft, Lange, Notowidigdo (2013)

- KLN – Probability of callback declines w/unemp duration.
- We find no relationship. – Why?
- Important differences between studies.
    - Outcome measure
        - KLN – Callback requesting an interview.
        - We use any callback.
    - Job Type
        - KLN – 3 job types (sales, customer service, admin/cler).
        - We use one type (low-mid level admin jobs).
    - Timing
        - KLN applications sent June 2011 - July 2012.
        - Our applications sent March 2012 - August 2014.
    - Choice of Cities
        - KLN use the 100 largest U.S. metropolitan areas.
        - We use 8 cities chosen to have high or low U.
    - Age of Workers
        - KLN study younger workers (ages 19-40, with 83% 23-31).
        - We study older workers (35-58).
- We use KLN's data to investigate the role of these differences.

Which Design Difference Could Account for Different Result?

- Outcome measure – Not clear.
- Job Type – This could matter.
  - One of the KLN job types (admin/clerical) matches nicely with the jobs applied for in our sample.
  - Our applicants are all female. KLN's applicants in this categories are 96.5% female.
  - We compare our and KLN's results for comparable category.
- Timing – Could matter.
  - KLN is earlier (mid-2011 to mid-2012) than FSvW (mid-2012 to 2014). KLN earlier in recovery.
- City Choice – Not likely to matter.
  - We compare results for KLN applications in our cities.
- Age of applicants – This could matter.
  - It may be that the longer employment history of older applicants makes current U duration uninformative.
  - We use KLN's data to investigate how their callback rates vary by age within their sample of young applicants.

General Approach to KLN-FSvW Comparison

- KLN graciously made their data available for reanalysis.
- We use these data to estimate a very simple logit model of the probability of callback.
    - Model includes *only* unemployment duration in months.
    - We report the marginal effect of unemployment duration on the probability of callback.
- We then repeat the same analysis using our data.
- We compare the results.
- We do this for various sample permutations that allow us to focus on particular possible explanations.

Outcome Measure

Sample	N Apps	Callback Rate	Marginal Effect U months
FSvW Data	12224	10.37	0.00001
			(0.00061)
KLN Callback/Interview	9236	4.54	-0.00086
			(0.00024)
KLN All Callback	9236	12.05	-0.00141
			(0.00024)

Note: Robust standard errors clustered by job id in parentheses.

- KLN all-callback rate (12.05%) is comparable to the FSvW callback rate (10.37%).

- KLN callback w/interview rate (4.54%) is much lower.

- Marginal effect of U duration on KLN all-callback rate smaller than effect on callback w/interview rate, but still statistically significant.

- Different outcome measure do not account for different findings.

- **Continue using comparable all-callback measure for KLN**

Analysis of the KLN Admin/Clerical Applications

Job Type / Sample	N Apps	Callback Rate	Marginal Effect U months
All KLN Jobs	9236	12.05	-0.00141
			(0.00024)
KLN Admin/Cler Jobs	2690	3.61	-0.00079
			(0.00037)
FSvW Admin Support	12224	10.37	0.00001
			(0.00061)

Note: Robust standard errors clustered by job id in parentheses.

- Much lower callback rate in KLN for the admin/clerical jobs.
- Significant negative relationship between duration and callback for KLN admin/clerical applications.
- => Job type does not account for difference in results.

Time Period: Analysis of the KLN and FSvW Applications

- Showed in previous slide that callback rate in KLN for admin/clerical applicants was much lower than that for the admin/clerical applicants in our sample.
- Could reflect that earlier in recovery labor market was weaker so employers could exercise more discretion of who to hire.
- Consistent with FSvW analysis showing smaller effects of observables where employer had high within-job callback rates.
- But this is suggestive, at best.

Analysis of the KLN Applications in the FSvW Cities

Group	N Apps	Callback Rate	Marginal Effect U months
KLN Non-FSvW Cities	8106	12.04	-0.00133
			(0.00037)
KLN FSvW cities	1130	12.12	-0.00192
			(0.00094)
KLN All Cities	9236	12.05	-0.00141
			(0.00024)
FSvW Data	12224	10.37	0.00001
			(0.00061)

Note: Robust standard errors clustered by job id in parentheses.

- Insignificant negative relationship between duration and callback in our cities (other than Portland ME).
- But not significantly different from effect in other cities.
- => City choice likely does not account for difference in results.

Analysis of Age in the KLN Sample

Age Category / Sample	N Apps	Callback Rate	Marginal Effect U months
KLN 19-22 Years Old	674	10.68	-0.00515
			(0.00186)
KLN 23-26 Years Old	3840	11.59	-0.00078
			(0.00054)
KLN 27-30 Years Old	3622	12.78	-0.00197
			(0.00055)
KLN 31-39 Years Old	1100	12.09	-0.00268
			(0.00099)
KLN All Ages	9236	12.05	-0.00141
			(0.00024)

Note: Marginal effects calculated from logit model of callback. Robust standard errors clustered by job id in parentheses.

- Marginal effect of duration varies by age (p-value $= 0.008$).
- $=>$ Perhaps the effect of duration on callback is not there for older workers (like those in FSvW).

Conclusions Regarding the Comparison of KLN and FSvW

- Choice of cities and job types (maybe) not important.
- Time period may be important.
  - KLN study is early in recovery, while FSvW is later.
  - Lower callback rate for admin/cler in KLN => Employers more selective.
  - Conclusion: Perhaps effect of duration on callback only relevant in very weak labor market.
  - But this is inconsistent with KLN finding that effect is larger (cross-sectionally) in strong labor markets.
- Age may be important.
  - KLN study has much younger applicants (99% $\leq$ 33) than FSvW study (all $\geq$ 35).
  - Within KLN study, the effect varies by age and strongest for their youngest workers (19-22).
  - Conclusion: Perhaps KLN result does not generalize to older workers (about whom the market has more information).
- General Lesson: Maybe limited external validity of experiments.

Outline of Talk

- Introduction
- Bayesian Updating Model
- Design of Audit Study
- Univariate Results
- Multivariate Results
- Why are findings different from Kroft-Lange-Notowidigdo?
- **Conclusion**

Discussion and Conclusion – Audit Study for Mature Workers

Examined the effect of unemployment duration, interim jobs and Age on callback rates

Three Main Findings:

1. No Effect of Unemployment Duration on Call Back.
   - Our null result on unemployment in contast to KLN, who focus on younger workers.
2. Being older lowers the callback rate substantially.
   - Consistent with discrimination against older workers.
3. Taking an interim job lowers callback rate substantially.
   - Employers appear to use interim job as signal.
4. Firms with higher needs appear to be less choosy.

Overall, findings consistent with view that employers solve signal-extraction problem, but for mature & older workers it does not involve unemployment duration.

Broader Methodological Implication – Lack of External Validity

- The findings of audit studies on the effect of unemployment duration on callback rates is not consistent.
  - Some (e.g., KLN) find strong negative effect.
  - Others (e.g., FSvW) find no significant effect.
- Our attempts to reconcile these effects are *ad hoc* and do not provide clear guidance on what is driving the differences in findings.
- Be very careful not to extrapolate from particular studies to other settings.
- These are isolated fragments of evidence that are suggestive rather than definitive.
- A suggestion: Use economics (modeling and theory) to understand (and even predict) heterogeneous effects.