Final Report: Are one-column or two-column resumes more successful?

Silas Gifford, Leo Le, Pavan Sabnis

Introduction

For our research project, we investigate whether single column resumes garner more attention from employers compared to dual column resumes. Resumes serve as a crucial initial point of contact for job seekers with potential employers, and the way employers perceive a resume can significantly impact interview invitations and job offers. Technological advancements including applicant tracking systems (ATS) have automated the initial resume screening process, adding complexity to resume design considerations. Our study seeks to determine the optimal resume format that is not only favorable for these technological systems but also for human recruiters. Existing research suggests that resumes with multiple columns may not be ATS-compatible and are generally less read by recruiters. Based on this, we hypothesize that using a single column resume format will result in higher positive response rates. We also expect that the technology sector will have fewer invitations to interview for two column resumes than any other sector because we believe that there is more ATS automation within that sector.

Experiment Details

Comparison of Potential Outcomes:

We created resumes tailored to three different sectors (technology, retail, and manufacturing) and formatted them twice, once with a single column and once with two columns. We submitted these resumes to job postings and tracked the number of invitations to interview after four weeks. We employed a within-subject design coupled with regression, where both types of resumes were submitted to the same job posting. This method effectively assigned both treatment (two column resume) and control (one column resume) to the same job posting. Our regression takes the following form:

$$Interview = intercept + a * column + b * sectorA + c * sectorB + d * job\_postingA + ... + z * job\_postingZ$$

This allows us to compare all one column resumes to all two column resumes with controls for sector and job posting/company. We anticipated that some companies might respond to a particular type of resume, while others might respond to both resumes regardless of format, and some might not respond at all. To account for this variation in responsiveness among companies, we incorporated regression fixed-effects by company into our analysis.

We also expect response rates to vary by sector and that different sectors may use ATS at different rates. These systems can play a significant role in screening and evaluating resumes before they are reviewed by hiring managers. Typically, ATS favor single-column resumes due to their straightforward formatting, posing a challenge for the two-column format. Addressing this challenge is complex, as it's uncertain which job applications will undergo processing through an ATS. However, we are committed to ensuring that both versions of the resume meet all other criteria for an ATS-friendly format. This includes ensuring conciseness, freedom from grammatical errors, and consistency in language usage.

To ensure the accuracy and fairness of our results, the treatment and control resumes were formatted similarly, with variations in experience descriptions and applicant names, while maintaining the same quality and years of work experience. We created two personas for each sector, and for each persona we formatted a one column and a two column resume. See Figure 1 for an example persona and the formatting of one vs two column resumes.

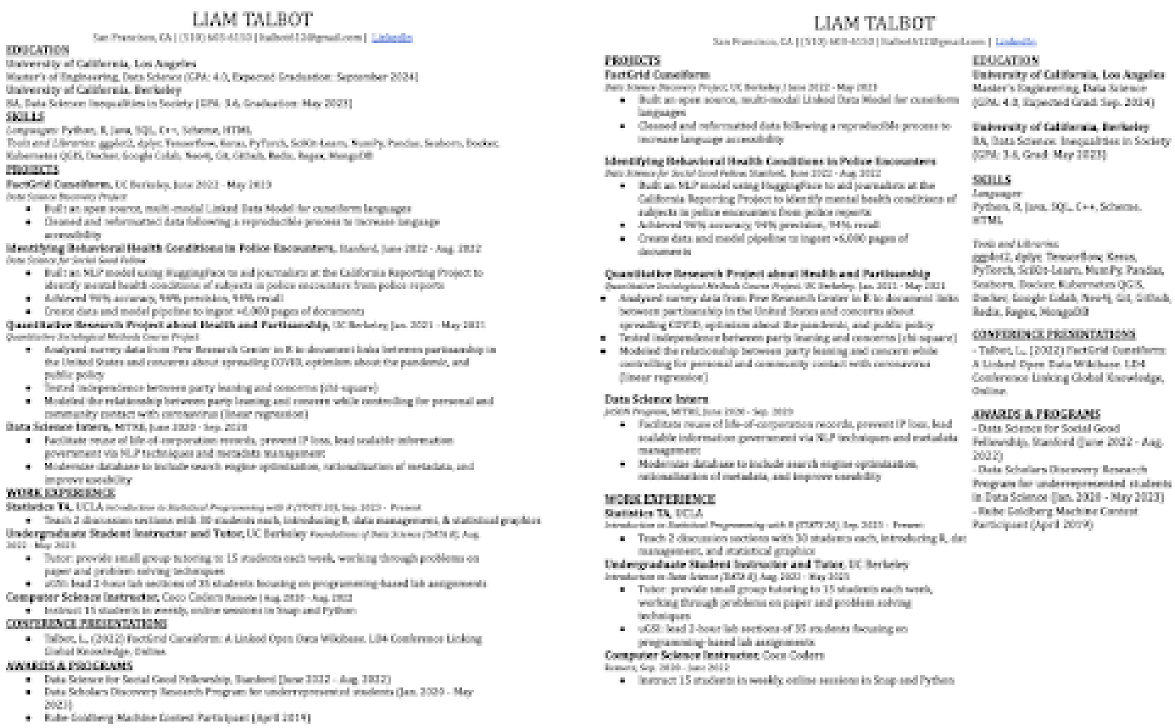


Figure 1: A pair of resumes for the persona of Liam Talbot in the technology sector

Randomization Process

We used a random number generator to determine which of the two personas submitted a one column resume and which of the personas submitted a two column resume to a given job posting for a sector. Specifically, if the generated number was 1, we would submit the one column version for persona A and the two column version of persona B. But if the number was 2, the opposite would occur with the one column version for persona B and the two column version for persona A being submitted. Assignment to persona A and persona B was fixed within a sector for the duration of the randomization process.

Each team member was responsible for applying to a minimum of 20 job postings for a sector. This goal results in 40 data points per sector and 120 data points in total.

Choosing Job Postings

We used LinkedIn to find all job listings. No postings older than 30 days were considered. We did not filter by location, salary, company, or application type such as “quick apply”. All job postings were given four weeks to respond once an application was submitted. After four weeks we recorded whether or not there was an invitation to interview. We consider a lack of response the same as a “no.”

For the technology sector, we looked at job titles related to data science, data engineering, and AI/ML. We applied filters for experience level: “internship”, “entry level”, and “associate”. Jobs were excluded if they required more than 3 years of experience because the applications were tailored for new graduate jobs. We submitted for 22 job postings in the technology sector.

For the manufacturing sector, we considered job titles of "Process Engineer" and managerial roles (ex: "Production Team Lead", "Manufacturing Supervisor", "Team Leader", "Shift Supervisor"). We filtered by experience, considering "associate" and "mid-senior level" positions. No jobs were excluded from this search and we submitted for 20 job postings in the manufacturing sector.

For the retail sector, we considered the following job titles: "Retail Associate", "Sales Associate", "Customer Service Representative", and "Retail Sales Consultant". We filtered by experience: "entry level" and "associate". Two job listings did not ask for a resume and were excluded from our search. Also excluded from the search were postings that required us to submit a social security number at the time of applying (N=4). In total, we submitted for 19 job postings for the retail sector.

Power Calculation

Our initial power analysis was for a simple t-test that did not take advantage of a within subject design and looked only at the difference in means for job postings that received a one column vs two column resume. Previous research on the topic suggests that the ATE may be small, as in Popham et al. (2016) who found no reliable effect for interview decisions between graphical resumes and non-graphical resumes, or large as suggested by Lookadoo & Moore's research (2024) that recommends against including columns, since including columns in a resume is generally not ATS-compatible, and therefore decreases the likelihood of the resume being selected. This simple analysis suggests that we'd want to submit over 400 resumes. See our power analysis submission for the code used to generate the following graph.

On the advice of David Reiley, we adjusted our design to allow for sector and employer fixed effects by using regression. The following analysis is a series of simple simulations that assumes each sector uses ATS at different rates and thus will have different ATEs. We have not designed the data in a way that would produce meaningful information about employers, so that control is excluded from this analysis. We discussed this approach with David Reiley in office hours.

```r
test_power <- function(
  mean_control = 0.5,
  mean_treat = 0.5,
  number_per_condition = 10) {

    d <- data.table()
    d[ , condition := rep(c('control', 'treatment'), each = number_per_condition * 3)]
    d[condition == 'control', sector := rep(c('a', 'b', 'c'), each = number_per_condition)]
    d[condition == 'treatment', sector := rep(c('a', 'b', 'c'), each = number_per_condition)]

    # ATS use in each sector: a > b > c
    # reasoning: control is one column, should not be affected by ATS use
    # more ATS use will lead to fewer successes for treatment group (two column)

    d[condition == 'control', outcome := sample(c(0, 1), size = number_per_condition * 3,
                             replace = TRUE, prob = c(1 - mean_control, mean_control))]

    sector_a_adv <- rnorm(1, mean = 0.1, sd = 0.05)
    mean_treat_a <- ifelse(mean_treat + sector_a_adv > 1, 1, mean_treat + sector_a_adv)
    d[condition == 'treatment' & sector == 'a', outcome := sample(c(0, 1),
                                                  size = number_per_condition,
                                                  replace = TRUE,
                                                  prob = c(1 - mean_treat_a, mean_treat_a))]

    d[condition == 'treatment' & sector == 'b', outcome := sample(c(0, 1),
                                                  size = number_per_condition,
                                                  replace = TRUE,
                                                  prob = c(1 - mean_treat, mean_treat))]

    sector_c_disadv <- rnorm(1, mean = -0.1, sd = 0.05)
    mean_treat_c <- ifelse(mean_treat + sector_c_disadv < 0, 0, mean_treat + sector_c_disadv)

    d[condition == 'treatment' & sector == 'c', outcome := sample(c(0, 1),
                                                  size = number_per_condition,
                                                  replace = TRUE,
                                                  prob = c(1 - mean_treat_c, mean_treat_c))]
    # regression
    mod <- d[, lm(outcome ~ condition + sector)]
    return(mod)

}
```

We tested sample sizes that consisted of 10, 15, and 20 job postings. Coefficients for each combination of sample size and estimated ATE are presented below. We consider a small ATE to be 0.05, a medium ATE to be 0.15, and a large ATE to be 0.3.

These results suggest to us that we should be able to detect a medium or large ATE for our chosen sample size of 20 job postings.

Table 1: Sample Size of 10

	Dependent variable:		
	outcome		
	Small ATE	Medium ATE	Large ATE
	(1)	(2)	(3)
2 column	0.100	−0.233*	−0.200*
	(0.131)	(0.124)	(0.118)
Sector B	−0.000	0.150	0.100
	(0.160)	(0.152)	(0.144)
Sector C	0.150	−0.200	0.200
	(0.160)	(0.152)	(0.144)
Intercept	0.450***	0.617***	0.700***
	(0.131)	(0.124)	(0.118)
Observations	60	60	60
R^2	0.030	0.137	0.079
Adjusted R^2	−0.022	0.091	0.030
Residual Std. Error (df = 56)	0.507	0.481	0.455
F Statistic (df = 3; 56)	0.583	2.959**	1.609
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2: Sample Size of 15

	Dependent variable:		
	outcome		
	Small ATE	Medium ATE	Large ATE
	(1)	(2)	(3)
2 column	−0.044	−0.022	−0.244**
	(0.106)	(0.104)	(0.100)
Sector B	−0.200	−0.233*	−0.200
	(0.130)	(0.127)	(0.122)
Sector C	−0.067	−0.200	−0.167
	(0.130)	(0.127)	(0.122)
Intercept	0.622***	0.744***	0.856***
	(0.106)	(0.104)	(0.100)
Observations	90	90	90
R^2	0.030	0.044	0.095
Adjusted R^2	−0.004	0.011	0.063
Residual Std. Error (df = 86)	0.504	0.492	0.474
F Statistic (df = 3; 86)	0.876	1.331	3.011**
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Sample Size of 20

	Dependent variable:		
	outcome		
	Small ATE	Medium ATE	Large ATE
	(1)	(2)	(3)
2 column	0.067	-0.200^{**}	-0.267^{***}
	(0.093)	(0.088)	(0.085)
Sector B	0.025	-0.050	-0.150
	(0.113)	(0.108)	(0.104)
Sector C	0.050	-0.225^{**}	-0.125
	(0.113)	(0.108)	(0.104)
Intercept	0.442^{***}	0.775^{***}	0.858^{***}
	(0.093)	(0.088)	(0.085)
Observations	120	120	120
R^2	0.006	0.079	0.095
Adjusted R^2	-0.020	0.056	0.072
Residual Std. Error (df = 116)	0.507	0.481	0.466
F Statistic (df = 3; 116)	0.238	3.336^{**}	4.063^{***}
Note:			*p<0.1; **p<0.05; ***p<0.01

Analysis

Data

The table below presents a segment of our data, encompassing six columns: Name, Sector, Salary, Columns, Result, and Job Posting.

- Name: Indicates the persona's name associated with the submitted resume.
- Sector: Categorizes the sector of the job application, with values including Manufacturing, Retail, and Technology
- Salary: Contains salary information; however, not all job postings included salary details, leading to its exclusion from our analysis.
- Columns: Specifies whether the resume format is single column (1) or double column (2).
- Result: Consists of binary values: 0 and 1. Here, 0 indicates a lack of response or a rejection, while 1 signifies that the employer has responded with further steps.
- Job Posting: Indicates the employer of the application.

For the purposes and scope of our experiment, we performed regression analysis with sector, columns, and job posting as covariates, aiming to explore their impacts on the outcome.

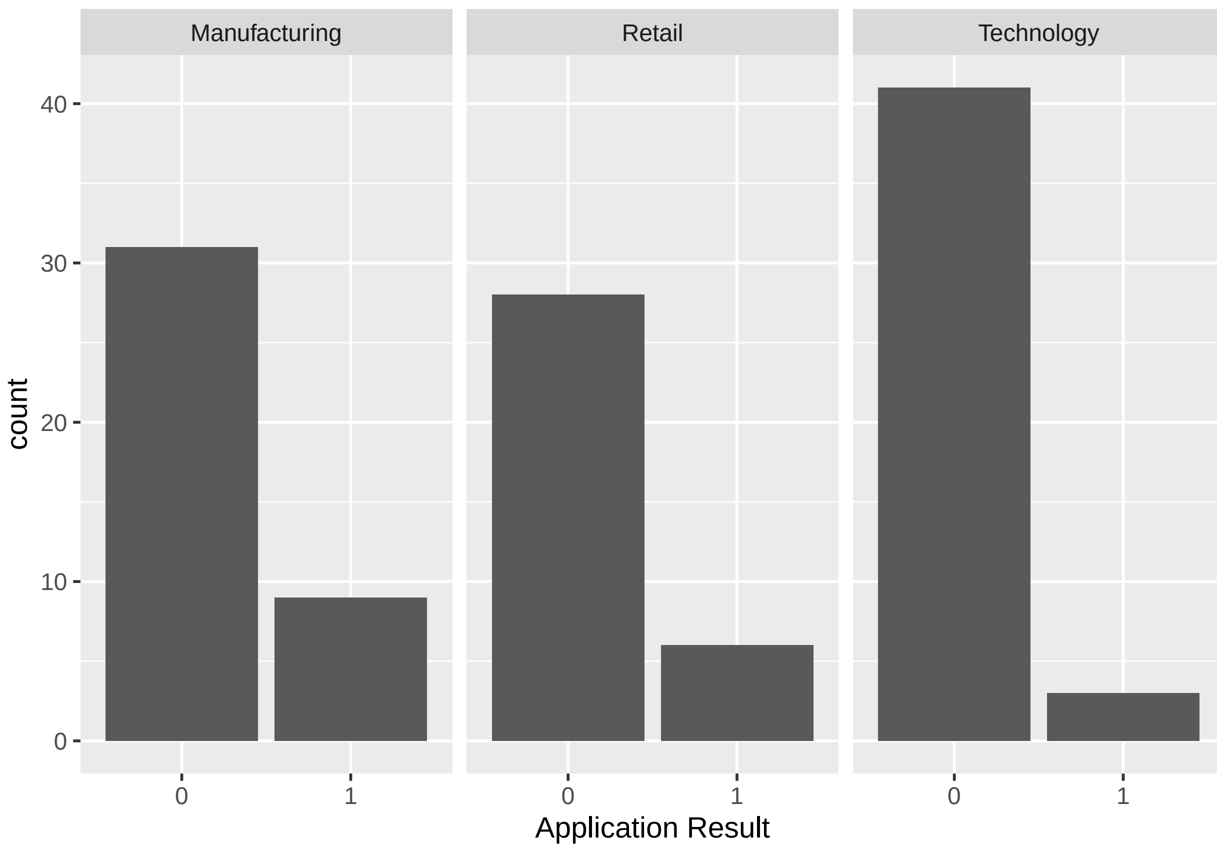
```r
d %>%
  head() %>%
  kable()
```

name	sector	salary	columns	result	job_posting
Nhan	Manufacturing	82000	1	1	Yakima Chief Hops
Nhan	Manufacturing	NA	1	1	The Kraft Group

name	sector	salary	columns	result	job_posting
Hoang	Manufacturing	NA	1	0	Packaging Corporation of America
Nhan	Manufacturing	NA	1	1	Ocean Spray
Nhan	Manufacturing	NA	1	0	Graphic Packaging International
Nhan	Manufacturing	NA	1	0	ND Paper

Here is how our outcome measure is distributed by sector.

```
d %>%
  ggplot(aes(x=as.factor(result))) +
  geom_bar() + facet_wrap(vars(sector)) +
  labs(x = "Application Result")
```



Models

To begin, we start by modeling the simple difference between one column and two column resumes without controls.

```
simple_model <- lm(result ~ columns, data = d)
simple_se <- sqrt(diag(vcovHC(simple_model)))
stargazer(simple_model, type = "latex", se = list(simple_se),
          header=FALSE, title = "Simple Model")
```

This model shows that changing from one column to two columns on a resume predicts a decreased outcome by -0.0338983. A 95% confidence interval for this coefficient is -0.1657362 to 0.0979396 so this result is not statistically significant.

Table 5: Simple Model

	Dependent variable:
	result
columns	−0.034
	(0.067)
Constant	0.203*
	(0.109)
Observations	118
R^2	0.002
Adjusted R^2	−0.006
Residual Std. Error	0.362 (df = 116)
F Statistic	0.258 (df = 1; 116)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
full_model <- lm(result ~ sector + columns + job_posting, data = d)
full_se <- sqrt(diag(vcovHC(full_model)))
stargazer(full_model, type = "latex", se = list(full_se),
          keep = c("sectorRetail", "sectorTechnology", "columns", "Constant"),
          header=FALSE, title = "Full Model")
```

Table 6: Full Model

	Dependent variable:
	result
sectorRetail	−0.000
	(0.983)
sectorTechnology	−0.500
	(0.695)
columns	−0.034
	(0.059)
Constant	0.551
	(0.701)
Observations	118
R^2	0.806
Adjusted R^2	0.608
Residual Std. Error	0.226 (df = 58)
F Statistic	4.073*** (df = 59; 58)
Note:	*p<0.1; **p<0.05; ***p<0.01

Including covariates increases the precision of this estimate. Our full model shows that changing the form of a resume from one column to two columns predicts a decreased outcome by -0.0338983 holding the job posting and sector constant. A 95% confidence interval for this coefficient is -0.1502978 to 0.0825012, again not statistically significant.

We have excluded the 58 coefficients for `job_posting` from this stargazer table for brevity and will do so in all stargazer tables.

The coefficients for the number of columns are not particularly large and they do not provide estimates statistically different from zero, but they may still have practical importance when submitting job applications en masse. We can also conclude that if you are in the technology sector, then you probably want to send out tons of applications because the positive response rate compared to the retail and manufacturing sectors is low.

Now we examine a model that includes an interaction term for sector and columns as discussed during the final presentation. However, we will first re-encode the `columns` variable to create a new variable: `treated`. As a reminder, `columns` is the number of columns on a resume, either `1` or `2`. `treated` will be an indicator variable that indicates whether a unit was treated or not: `0` for a one column resume and `1` for a two column resume. This re-encoding simplifies the interpretation of interaction terms.

```
d <- d %>%
  mutate(treated = as.numeric(columns == 2),
         tech = as.numeric(sector == "Technology"))
```

We also made an indicator variable for the technology sector, `tech`, in order to investigate the difference between the technology sector and non-technology sectors which, as we noted above, have quite different positive response rates. We create a linear regression with `tech` (as opposed to `sector`) here.

```
model_tech_interact <- lm(result ~ tech + treated + job_posting + tech*treated, data = d)
tech_interact_se <- sqrt(diag(vcovHC(model_tech_interact)))
stargazer(model_tech_interact, type = "latex", se = list(tech_interact_se),
          keep = c("treated", "tech",
                   "tech:treated", "Constant"),
          header=FALSE, title = "Model with Interaction for Technology")
```

As in the previous model, we see a large negative coefficient corresponding to the technology sector, -0.5632678, due to the low positive response rate. The coefficient for `treated` is more than twice as large in magnitude than the coefficient we saw above for the corresponding coefficient for `columns`, but it still is not significant. The interaction term is positive, suggesting that two column resumes actually do better in the technology sector than one column resumes. This makes sense given our data; there were three positive responses within the technology sector and two of the three positive responses were for two column resumes. Again, this result is not significant so if this experiment was replicated we make no claims that the same effect would be found, or even an effect with the same sign. A larger sample size and/or a higher positive response rate for the technology sector would aid our analysis, but alas.

```
d %>%
  filter(result == 1 & sector == "Technology") %>%
  kable()
```

name	sector	salary	columns	result	job_posting	treated	tech
Silas	Technology	NA	1	1	Bose	0	1
Liam	Technology	NA	2	1	Bose	1	1
Silas	Technology	80-146k	2	1	Harvard Business School	1	1

For completeness, we include the following model that interacts `sector` with `treated`.

```
model_interact_full <- lm(result ~ sector + treated + job_posting + sector*treated, data = d)
full_int_se <- sqrt(diag(vcovHC(model_interact_full)))
```

Table 7: Model with Interaction for Technology

	Dependent variable:
	result
tech	−0.563
	(0.671)
treated	−0.081
	(0.086)
tech:treated	0.127
	(0.108)
Constant	0.541
	(0.669)
Observations	118
R^2	0.813
Adjusted R^2	0.616
Residual Std. Error	0.224 (df = 57)
F Statistic	4.125*** (df = 60; 57)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
stargazer(model_interact_full, type = "latex", se = list(full_int_se),
          keep = c("treated", "sectorRetail", "sectorTechnology",
                   "sectorTechnology:treated", "sectorRetail:treated", "Constant"),
          header=FALSE, title = "Full Model with Interactions for All Sectors")
```

This model suggests that a two column resume performs slightly worse in the retail sector (interaction term for the retail sector and treated: -0.0676471) and slightly better in the technology sector (interaction term for the technology sector and treated: 0.0954545) compared to two column resumes in the manufacturing sector. These results are not statistically significant so we cannot reject the null hypothesis that these coefficients are truly different from zero. Our hypothesis that the technology sector would have fewer positive responses for a two column resume is still up for question, although this results appear to contradict it.

In all of our regressions, we found a small (<10%) and statistically insignificant effect when submitting a two column resume (as opposed to a single column resume).

Table 9: Full Model with Interactions for All Sectors

	Dependent variable:
	result
sectorRetail	0.034
	(0.974)
sectorTechnology	−0.548
	(0.710)
treated	−0.050
	(0.073)
sectorRetail:treated	−0.068
	(0.186)
sectorTechnology:treated	0.095
	(0.098)
Constant	0.525
	(0.708)
Observations	118
R^2	0.814
Adjusted R^2	0.612
Residual Std. Error	0.225 (df = 56)
F Statistic	4.022*** (df = 61; 56)
Note:	*p<0.1; **p<0.05; ***p<0.01