Methodology: Monitoring trends in remote and/or hybrid jobs with Indeed data

Shawn Tham and Chris Glynn
June 28, 2023

Indeed Hiring Lab previously released the proprietary *Remote Job Tracker*, measuring the aggregate share of job postings in a specific location that included keywords associated specifically with remote work. We start by searching if the job location is remote before doing a keyword search on the job posting text to identify remote/hybrid jobs. Today, with hybrid work emerging as a primary modality of flexible work — which itself is a topic of growing interest to job seekers, employers, and policymakers alike — we are updating and renaming the Remote Tracker to the Remote & Hybrid Job Tracker. The methodology behind this re-released series expands upon the definition of remote work previously used to construct the Remote Job Tracker, and now includes terms more generally associated with hybrid and flexible work arrangements.

Methodology

Denote the ith job posting in location l at time t by $JP_{i,l,t}$, and let $N_{l,t}$ be the total number of postings in the same location and time. For each job posting, we sum across an indicator associated with every word, denoted by w, in $JP_{i,l,t}$, effectively counting the number of words in the job posting that are also in the hybrid and remote keyword dictionary, D_t. When at least one word in the posting is included in dictionary D_t, posting $JP_{i,l,t}$ is counted as a remote or hybrid posting.

Note that the dictionary itself is indexed by time. For all dates through December 2020, the previous dictionary with terms focused on remote work is utilized. Starting in January 2021, an expanded dictionary is used that combines the existing keyword set and additional hybrid terms and key phrases.

$$RH_{\ell,t} = \sum_{i=1}^{N_{\ell,t}} \mathbb{1}\left\{\left(\sum_{w \in JP_{i,\ell,t}} \mathbb{1}_{w \in D_t}\right) > 0\right\}$$

$$\rho_{\ell,t} = \frac{RH_{\ell,t}}{N_{\ell,t}} \times 100$$

January 2021 was when standalone hybrid job postings emerged in our data, according to an internal analysis. To prevent false positives prior to January 2021, only the previous dictionary of remote terms is used.

Case Study: US (2020 till April 2023)

Percentage (%) of job postings containing remote and/or hybrid terms

7-day Moving Average
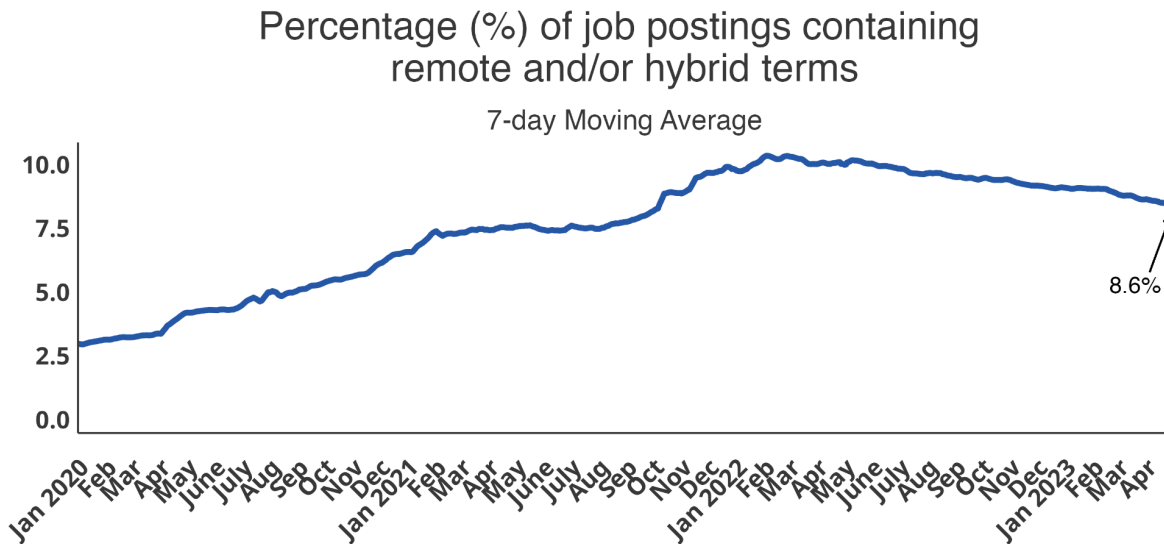
Source: Indeed

indeed

Line graph titled 'Percentage of job postings containing remote and/or hybrid terms' with a vertical axis from 0% to 10%. The graph covers from January 2020 to April 2023. It shows remote job posting numbers rising quickly through most of 2020 and 2021 before peaking in March 2022 and declining gradually

The share of jobs classified as remote peaked in early 2022, with a small but consistent decline into 2023. The relative increase in the share of U.S. jobs classified as hybrid or remote after January 2021 is between 0.5 - 0.7%.

Type 1 and Type 2 Errors

Our keyword search approach has limitations, most notably false positive and false negative rates. Consider the word *hybrid* itself, for example, which may refer to *hybrid work* but also to *hybrid cars*, *hybrid clouds*, or *hybrid education*. Keyword searches induce a certain share of postings that are flagged as remote or hybrid, but are in fact unrelated to flexible work arrangements.

Similarly, there are remote and/or hybrid postings that our keyword search does not pick up, resulting in a false negative. To control these false positive and false negative rates, we constructed numerous random samples of job postings previously classified as remote or hybrid

using a given keyword list, and calculated the sample's false positive and negative rates. We then adjusted the keyword list to minimize type 1 and type 2 error rates. Given the manual nature of inspecting large numbers of job postings to estimate type 1 and type 2 errors, a formal bootstrap procedure for estimating standard errors on $\rho_{l,t}$ is not practical.

False Positive Example (Link [here](#))	False Negative Example (Link [here](#))
Your chance to join a world renowned and respected brand in a newly created hybrid role. The emphasis is on understanding business as well as being able to manipulate digital data into useable insights for the advertising and marketing and ecommerce teams. Amazing company culture and vibe and terrific opportunity for personal development within the role. Generously salaried to £45K + Bens.	This is a full time permanent position with benefits Hourly pay of 30-35+ per hour 401K with 3% match 3 Weeks PTO Life insurance Health and Dental benefits Must be in office 4 days per week

The above images are screenshots of job postings from Indeed website that shows an example of a false positive and false negative remote/hybrid job posting when using the keyword "hybrid role" to search.

In the false positive example, the keyword "hybrid role" is included in the dictionary, but hybrid seems to refer to the role itself being a hybrid between business insight and digital data. In the false negative example, there is no mention of hybrid or remote work, but it clearly states that the person must be in the office 4 days per week, which is not captured in our dictionary.

Global Coverage

Once this dictionary of keywords (and phrases) was identified for English-language job postings in the U.S., Canada, the United Kingdom, and Australia, we worked with linguists to effectively translate the keyword list into Spanish, French, German, and other languages. We relied on local economists and native speakers to validate the translations and add language-specific terms not included in the original dictionary.

Relative Trends in Hybrid Work

It is not possible to easily identify jobs as exclusively hybrid, exclusively remote, or both hybrid *and* remote. Estimating the share of postings that is exclusively hybrid, for example, would require a keyword list with a designation of each keyword as remote, hybrid, or both. But where would the phrase "flexible work" fit in? And even if each keyword/phrase did have a designation, many job postings contain keywords that would be associated with both hybrid and remote terms. Cleanly identifying the compositional share of the overall remote/hybrid universe is not possible.

However, we can track the share of jobs captured by only those terms added in January 2021, denoted by the set difference in the dictionaries $(D_{t_1} \setminus D_{t_0})$ where $t_1 \geq January$ 2021 and $t_0 \leq December$ 2020.

This does not allow us to cleanly identify hybrid-only jobs, as there are postings with hybrid terms that also include terms associated with remote work. Nonetheless, our method does allow us to quantify the relative overlap in jobs that include both hybrid terms and remote terms — allowing us to observe an increasing trend in work flexibility from positions that increasingly feature words associated with hybrid work (essentially the Hybrid & Remote and Hybrid components of the Venn diagram below).
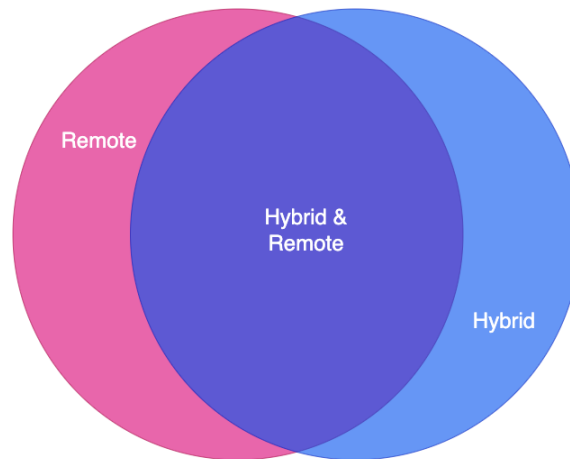
*The above image demonstrate the overlapping nature of remote and hybrid job postings.*

Effectively, this metric is the share of job postings classified by our method as hybrid or remote (all sets in the Venn diagram above) that include terms in the dictionary added in January 2021 (the overlap and hybrid-only sections of the Venn diagrams), denoted formulaically by:

$$\frac{\rho_{l,t} \mid \left(w \in D_{t_1} \setminus D_{t_0} \right)}{\rho_{l,t}} .$$

While this doesn't specifically capture the share of postings that are hybrid, it does allow us to roughly quantify the relative share that also includes hybrid terms in our dictionary post-January 2021.

Remote/Hybrid Job Postings Breakdown
US
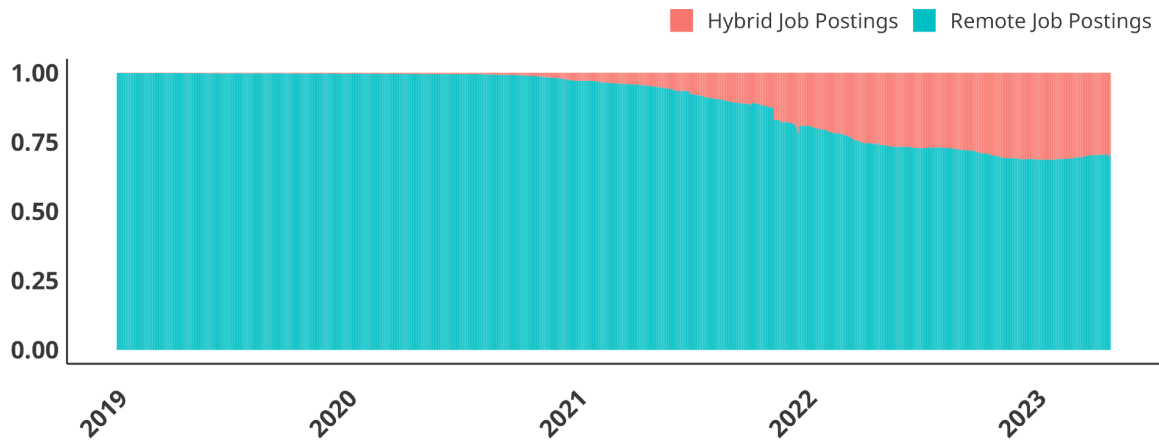
Source: Indeed

Line graph titled 'Remote/Hybrid Job Postings Breakdown' with a vertical axis from 0 to 1. The graph covers from January 2019 to April 2023. It shows the breakdown of job postings being pulled using remote only and hybrid only keywords. "Hybrid job postings" in the graph above refers to job postings added using the new hybrid keywords. Remote job postings in the graph above refers to job postings captured using old remote keywords. Because of the limitation in identifying fully remote and hybrid job postings accurately, it does not reflect the exact breakdown. Job postings pulled using hybrid keywords only started to form a bulk of the numbers from 2021 onwards.

Despite the inability to exactly capture remote and hybrid job postings, the Indeed Remote & Hybrid Tracker does accurately represent the growth trends of remote and hybrid jobs in the labor market. Given the growing interest in remote jobs, these temporal changes in remote and hybrid job postings help to identify changes in the labor market and demonstrate how fast employers are responding through new remote and hybrid working styles catering to jobseekers.

Appendix A
Table of Old and New Remote/Hybrid Keywords

Old Remote Keywords	New Hybrid Keywords
1. work life balance 2. flexible working 3. flexible work 4. flexible working	1. #LI-hybrid 2. combination of in-office 3. Flexible Hybrid 4. Flexible/Hybrid

arrangements
5. flexitime
6. work-life balance
7. flexible hour
8. flexible hours
9. flexible working hour
10. flexible schedule
11. flexible working hours
12. staggered hours
13. compressed hours
14. annualized hours
15. flexible scheduling
16. flexible work schedules
17. flexible working
      arrangement
18. work remotely
19. work from home
20. working remotely
21. working from home
22. remote work teleworking
23. remote work
24. teleworking
25. telework
26. telecommute
27. telecommuting
28. work at home
29. remote
	5. hybrid in-office
	6. hybrid office
	7. Hybrid role
	8. hybrid roles
	9. hybrid schedule
	10. Hybrid split
	11. hybrid telecommute
	12. hybrid Telecommuting
	13. Hybrid Telework
	14. hybrid teleworking
	15. hybrid work
	16. hybrid workforce
	17. hybrid working
	18. hybrid workplace
	19. Hybrid/Flexible
	20. Hybrid/Remote
	21. LI-Remote
	22. Location: Hybrid
	23. Remote/Hybrid
	24. work from the office or a
	      hybrid of time spent at
	      both

Appendix A: Table that contains old and new keywords used to identify remote and hybrid job postings

Table that contains old and new keywords used to identify remote and hybrid job postings.