Working from home: estimations without surveys[*]

true   true   true

November 22, 2020

Abstract

The COVID-19 crisis has forced great societal changes, including forcing many to work remotely (work from home) in an effort to increase social distancing. The ability to work from home has long been considered a perk, but we have few estimates of how many are actually able to work from home. Social media has been quick to argue that the people who are able to work at home are already privileged, highly educated and highly paid, while those who still have to go to work are more often in low-paid but critical roles such as drivers and grocers. This paper attempts to estimate the share of the Norwegian workforce able to work remotely by combining register-based labor statistics, annotated job ads and the official documentation of the ISCO-08 standard which is used both in labor statistics and in the job ads. We find that approximately 36 % of Norwegian jobs can be performed at home.

Introduction

Covid-19 pandemic hit the world hard and unprepared. In a study of the Spanish flu, (**?**) shows that non-pharmaceutical interventions known as "social distancing" during a pandemic can significantly reduce the disease transmission and lower both the peak and cumulative excess mortality. Learning from the historical lessons, many countries, including Norway, implemented measures to limit physical contacts between people. Encouraging working remotely (working from home) is one important part of these measures. However, concerns are raised quickly among experts and ordinary people on both its effectiveness and impact on social fairness. They argue that such policy will help only those who are high educated, highly paid and have comfortable jobs. The people who are much more susceptible benefit little from such measures. These concerns are well founded, but they ignored the externality of this measure. In fact, the measure will reduce the risk of infection indirectly for those who cannot work from home as well, since it will reduce both the potential infection risk and frequency of exposures to such risk. Here, the prevalence of remote feasible jobs matters. If there is only a very small fraction of working force can work from home, the effect may be ignorable to people others than those who work from home. Large prevalence of such jobs will imply higher effectiveness and lower negative consequence on social equality. However, there is very limited knowledge on prevalence of remote feasible jobs. In this paper, we try to answer the question: who and how many can work from home in Norway. Our analysis is based on three different types of data: a) the information of tasks described in the ISCO-08 standard (**?**). b) Job advertisements published by the Norwegian welfare administration (NAV) between January 2012 and march 2019, in which there are mentions of remote possibilities. c) official statistics on occupational employment published by Statistics Norway. The feasibility of

---

[*]TBD

an occupation is evaluated using ISCO descriptions of tasks to be performed. Detailed explanation of such classification is explained in the following in section sec:methodology [!TODO]. The classifications are then linked to the job advertisement data for consistency check and potential bias connection. Combing the feasibility classifications with occupational employment statistics, we then obtain information on the prevalence of remote feasible jobs and some characteristics of workers with such jobs. With help of the employment register, this can be done not only on national level but also on smaller geographical areas: for example, on municipalities.

We find that approximately 37 % of Norwegian jobs can be performed at home. As we expected, remote friendly jobs are often paid better than non-remote friendly jobs. There are large geographical differences in the prevalence of remote friendly jobs. There are larger share of remote friendly jobs in urban areas than rural areas. To some extent, this is a good news given that urban areas often facing a large challenges in containing the spread of Covid-19 given their high population density.

The method we applied here is similar to a very recent study by (**?**). They explores the same question in the United States using the Occupational Information Network (O*net) surveys covering "work context" and "generalized work activities". Unlike their study, we reply on much less conventional types of data: the description of tasks and contexts from job advertisements, which in turn requires new techniques for information extraction: manual labelling and natural language processing (NLP) techniques. Our study shows that combining conventional and unconventional data sources can lead to novel information that are useful for both official statistics, research and social policy making.

There were two surveys in norway in which some remote work feasibility is asked: The Norwegian labor force survey and a recent survey by the Norwegian Institute of Transport Economics(TØI). The Norwegian labor force survey has asked about remote-possibilities, covered in a report by (**?**). The question was whether the respondent had the opportunity to work from home at times, which is not to say that the job could be performed remotely in its entirety, and neither to say that those who weren't given the opportunity couldn't have worked from home if they had the option. The survey by TØI is designed specially for the Covid-19 situation with the main focus on the effective of remote work (**?**). While the survey does provide an overall estimate of the prevalence of remote jobs, information asked is rather limited. Neither occupational nor geographical aspects of the jobs are collected. Nevertheless the results we get were broadly similar to these surveys, while our analysis provides much more detailed information.

Method

The International Standard Classification

The ISCO standard organizes jobs into a set of groups according to the tasks and duties undertaken in the job. Using the detailed task descriptions listed in the ISCO-08 documentation, we try to provide a assertion of whether an occupation is likely able to be performed from home. To do this, we created a public labeling job through Amazon Mechanical Turk (**?**). Each occupation was presented together with a brief description. The exact question formulation was "Can this type of job likely be performed from a home office?", and an example of a job description could be:

> Social work and counselling professionals provide advice and guidance to individuals, families, groups, communities and organizations in response to social and per-

sonal difficulties. They assist clients to develop skills and access resources and support services needed to respond to issues arising from unemployment, poverty, disability, addiction, criminal and delinquent behaviour, marital and other problems.

The respondent was asked to evaluate wether it was likely that the job could be performed primarily from a private home. The alternatives were "Yes", "No" and "Unknown", which were provided with the following description:

1. *Yes: This job can be performed primarily from an office in a private home*

2. *No: Substatantial parts of this job must be performed outside the employees home*

3. *Unknown: There is not enough information to decide*

In order to reduce the serendipity in the labels, we acquired five labels from different respondents for each occupation, and we provided an `uncertain` option in addition to the `yes/no` options in order to reduce arbitrary responses to uninformative occupation descriptions. The final labels include an uncertainty measure which shows that some of the occupations were evaluated differently by different annotators, but no occupation was given a final label of "Unknown" which means we can treat the remote-friendly annotation as a binary variable.

Since the job was on Mechanical Turk, there respondents were not subject matter experts, and likely reside in different countries. This adds to the importance of obtaining more than one label per occupation, but the number of labels does not correct for possible cultural differences - it is possible that some jobs that cannot be performed remotely in other countries can be performed remotely in Norway. We should consider the annotations as `international`, which is also true for the ISCO-08 standard itself.

The initial results showed that for 77% of occupations, there were at least four concurrences. The remaining approximately 23% of occupations were more contested. There may be several reasons for this lack of quorum, one welcome and two unwelcome:

1. There are real differences in opinion as to wether a job can be done from home.
2. Workers are opportunistic, answering at random to minimize time spent on each task.
3. The information provided was insufficient, leaving workers to fill in the blanks themselves.

Real differences in judgement may be informative and increase our measure accuracy by incorporating several experiences and views on remote feasibility. The two other reasons for lack of quorum can be combatted by collecting more labels and by providing a better description of a job.

The occupations with a clear lack of quorum are annotated twice more, using two different MTurk panels. One panel of 10 workers who all are "masters" according to MTurk (a simple albeit vague checkbox indicating a high acceptance rate), and a panel of 5 workers who in addition to being "masters" have an task acceptance rate above 80 % (meaning that at least 80 % of their work has been accepted by other requesters). The threshold of 80 % was chosen mostly at random, but somewhat informed by comments on mturk user boards. In addition, one of the authors made separate annotation on the dataset.

The explanatory text on these rounds of annotation included the full description of the occupation, including descriptions of actual tasks. Due to the increase in text to read, the reward for workers was increased.

Neither of the rounds of MTurk annotation yielded any significant level of consensus among the workers - the consensus was only marginally better than what we could have expected from a

bootstrap resampling of the original data.

By carefully reviewing and annotating the selected occupation descriptions, some causes of ambiguity stands out: 1. Some jobs may or may not be performed from home depending on the home. Sowing, woodworking and similar types of craft can plausibly be done from home provided the home is spacious and properly equipped. On a longer time-horizon this may be plausible to many, while few may be able to bring this type of work home on short notice. 2. Some jobs can be done from home for a short while, postponing in-person meetings until a later time. 3. Some jobs can be done from home, but with a lower quality result. A perhaps poignant example are teachers, who have shown an ability to teach via video even in lower elementary education, but few parents would agree that this is an acceptable long-term or even medium-term solution. 4. Some jobs are very dependent on the technology in use at an employer. Filing clerks may have to stay in the office if filing is mostly on paper, but as more and more documents are digital only, such jobs may be done from home. The same is true for certain types of systems administrators in IT: If they are responsible for on-premise datacenters they may have to be physically present, while if they are using cloud-computing there are no physical servers to access.

Naturally, noise is still a factor: One may ponder how in all three rounds of MTurk annotation at least one worker submitted that a prison guard could work from home.

Lastly, a note on layman vs expert opinions. It is obvious that the workers performing the classifications are not labor market subject matter experts, and so the results are not authorative in reflecting the intention and original meaning from the creators of ISCO. For our purposes, this may only be a minor inconvenience, as we compare the annotations to occupational labels largely assigned by employers around Norway as part of their monthly tax reporting.

Consistency check using the job announcements data

In order to evaluate the annotations from Mechanical Turk, we use job advertisements from the Norwegian welfare administration (NAV). The job advertisements have been published as open data by NAV, and contain the text, title, employer information, and annotations made by subject matter experts at NAV including the occupational code (ISCO) of the job. The dataset covers January 2002 through march 2019 (a total of 2.6 million ads), but due to changes in the ISCO structure (a switch from ISCO-98 to ISCO-08 in 2012), we only use ads from 2012 and onward. Furthermore, the volume of NAV ads increased sharply in 2018 due to new sources of data. From 2018 the NAV data is close to complete in covering formally advertised jobs in Norway. Some shops still only advertise in their shop window, but the vast majority of jobs posted online are now also posted on NAV.

Because the possibility to work from home is a perk for many, some employers mention it in their job ads in order to attract candidates. We search the texts for mentions of `hjemmekontor` and `heimekontor`, two distinctive words unlikely to mean anything other than the possibility of working from home. Since far from every employer advertise this possibility, it is difficult to say anything about the total number of remote-friendly jobs from these ads. It may however say something important about the relative frequency of remote-friendly jobs across broader occupational groups, which we can use to validate the results from Mechanical Turk.

We have also tried to combine the two data sources in a more rigorous way to correct possible bias in the annotation labels. However, this can only be done under rather strong assumptions on employee's behavior. Detailed discussion can be found in the appendix.

Evaluate Remote feasibility

With the estimated remote feasibility for each occupation, we use the Norwegian labor market data from Statistics Norway (SSB) to evaluate the remote feasibility in Norway. The data from SSB covers employment, earnings and demographics. Employment data is register-based statistics which comes from two tables, covering the number of employed in Norway, divided by municipality and occupation respectively. Earnings data is also register based, and shows average monthly earnings by occupation. The demographics data shows population and density per municipality.

A note on nomenclature: For brevity, we sometimes refer to "remote" occupations rather than "occupations that can be performed remotely". We use the terms interchangeably, always referring to occupations that can be performed from home. This does not mean that such employees in actuality work from home either permanently or occasionally.

Remote feasibility using the annotation data

From the Mechanical Turk annotations, around 28 per cent of the occupations can likely be performed from home. Combining the annotations with labor statistics published per occupation, we find that 875 344 wage employers, 36 per cent of the workforce, are likely able to work from home.

From the Mechanical Turk annotations we see that 34 per cent of the occupations can likely be performed from home. Taken together with the labor statistics per occupation, we find that 939 898 jobs, 37 per cent of the workforce, are likely able to work from home.

From the Mechanical Turk annotations, around 34 per cent of the occupations can likely be performed from home. Taken together with the labor statistics per occupation, we find that 939 898 wage employees, 37 per cent of the workforce, are likely able to work from home.

Splitting the annotated data into occupational groups, we estimate what percentage of occupations are remote-friendly across occupational groups. The results are presented in Table tab:Percent_occupation} [!TODO]. For each broad occupational group, the share of jobs that can be performed remotely varies from 0 to 67 per cent. Academics and managers are both groups where more than half of the employees can be done remotely, but an even higher share of clerical support workers are likely able to do their jobs from home.

```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Number of jobs} & \\
##  \cmidrule(lr){2-3}
## Occupational Group & Total & Remote-friendly & Percent remote friendly \\
## \midrule
## Managers & $222 678$ & $146 344$ & $65,7\%$ \\
## Academics & $652 356$ & $330 783$ & $50,7\%$ \\
## Technicians and associate professionals & $374 858$ & $159 296$ & $42,5\%$ \\
## clerical support workers & $169 230$ & $106 668$ & $63,0\%$ \\
## Service and sales workers & $573 415$ & $153 136$ & $26,7\%$ \\
## Skilled agricultural, forestry and fishery workers & $21 631$ & $3 671$ & $17,0\%$ \\
## Craft and related trades workers & $219 843$ & $26 367$ & $12,0\%$ \\
## Plant and machine operators and assemblers & $163 197$ & $11 374$ & $7,0\%$ \\
```

```
## Elementary Occupations & $134 400$ & $2 259$ & $1,7\%$ \\
## \bottomrule
## \end{longtable}
```

In general, occupations that can be performed remotely also pay better, as shown in table tab:Earning [TODO]. The same pattern is also found when we split the data by occupational group (Figure figure:earning [TODO]). However, we see two exceptions: for academic professions where a lot of the remote-friendly occupations occur, the average wage for non-remote workers is slightly higher. And the distribution of wages have a much larger spread as well. The wage differential among remote and non-remote Craft and related trades workers is much larger.

```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrrr}
## \toprule
## Remote probability & \# of ocupations & \# of jobs & Average earnings & Median earnings \\
## \midrule
## High (>0.8) & $50$ & $406 745$ & $55 374$ & $48 754$ \\
## Medium (0.2 - 0.8) & $108$ & $877 310$ & $45 300$ & $41 867$ \\
## Low (<0.2) & $205$ & $1 499 120$ & $43 384$ & $40 973$ \\
## \bottomrule
## \end{longtable}
```

We simplify the remote probability further for the moment by converting it to a binary classification, simply rounding the probability to an integer (0 or 1). This allows us to present a more comprehensible view of the wage differences between remote- and non-remote jobs.
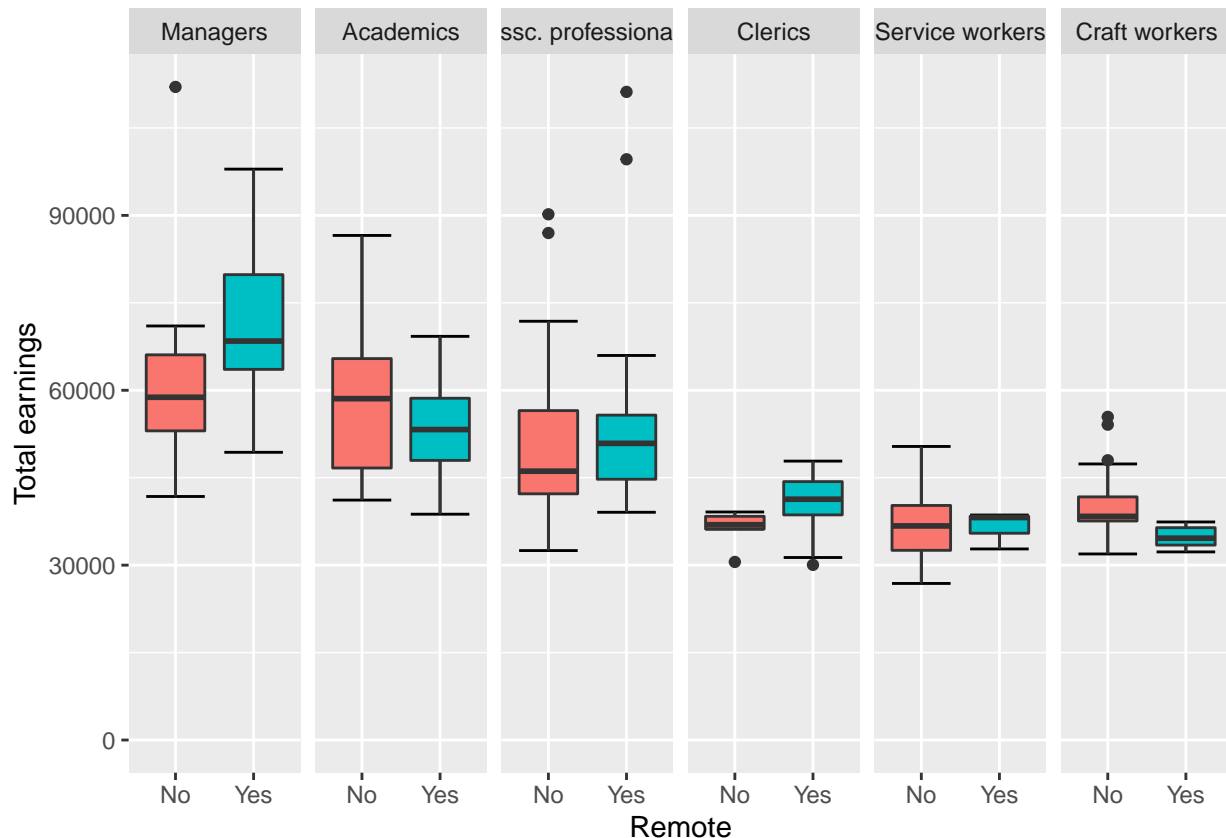
Variation of prevalence of remote feasibility across different regions

The geographic location of jobs have been a point of interest for years, amid both pressure for workers to centralize and specialize, and fears of increased inequality between cities and rural areas. Figure figure:geo[TODO] shows the percentage of workers can work from home in norway. As we expected, cities have a higher share of remote-friendly jobs, which may be fortunate given the need for social distancing. The pattern looks clear, especially in the area surrounding Oslo but also the other cities like Bergen, Trondheim and Stavanger seems to stand out on the map.
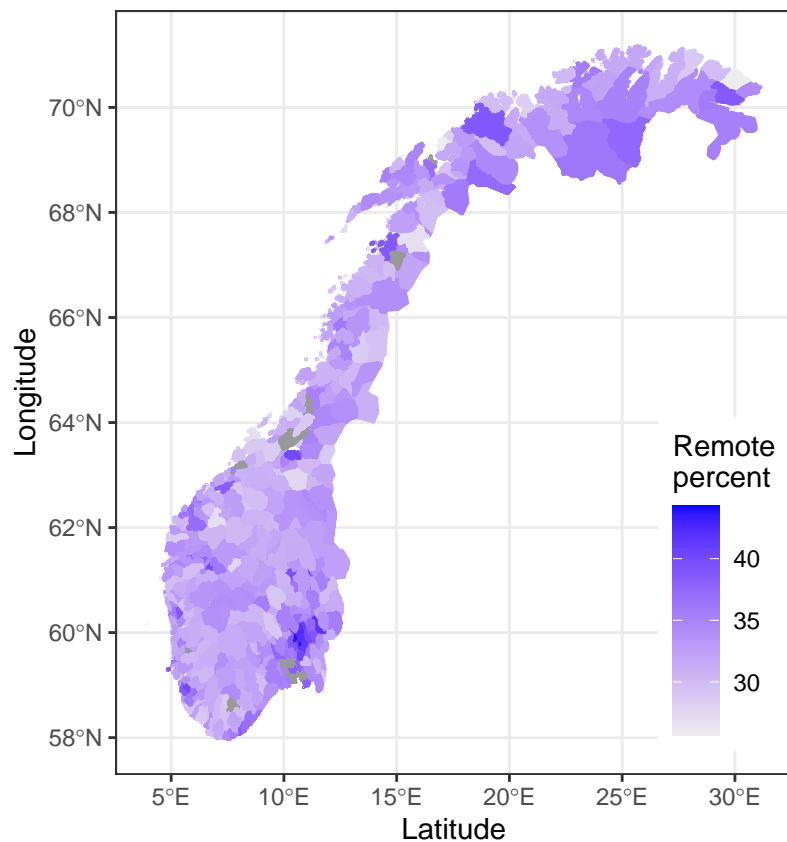


```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Number of jobs} & \\
##   \cmidrule(lr){2-3}
## Municipality & Total & Remote-friendly & Percent remote friendly \\
## \midrule
## Oppegård & $12 964$ & $5 743$ & $44,3\%$ \\
## Bærum & $60 596$ & $26 578$ & $43,9\%$ \\
## Asker & $28 960$ & $12 544$ & $43,3\%$ \\
## Oslo & $344 149$ & $145 368$ & $42,2\%$ \\
## Frogn & $7 308$ & $3 051$ & $41,8\%$ \\
## \bottomrule
## \end{longtable}

## \captionsetup[table]{labelformat=empty,skip=1pt}
```

```
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Number of jobs} & \\
##  \cmidrule(lr){2-3}
## Municipality & Total & Remote-friendly & Percent remote friendly \\
## \midrule
## Båtsfjord & $1 023$ & $263$ & $25,7\%$ \\
## Torsken & $373$ & $96$ & $25,7\%$ \\
## Sørfold & $842$ & $226$ & $26,9\%$ \\
## Stranda & $2 318$ & $625$ & $27,0\%$ \\
## Frøya & $2 494$ & $673$ & $27,0\%$ \\
## \bottomrule
## \end{longtable}
```

Table tab:5 most [TODO] and table tab:5 least[TODO] list the top and bottom 5 municipalities in terms of remote friendly jobs in Norway. We see that there are large heterogenity across different regions. Interpolating from the annotations we estimate that 42 percent of the jobs in Oslo can be done from home. On the other end of the spectrum, in municipalities like Båtsfjord less than a quarter of the jobs can be done remotely.
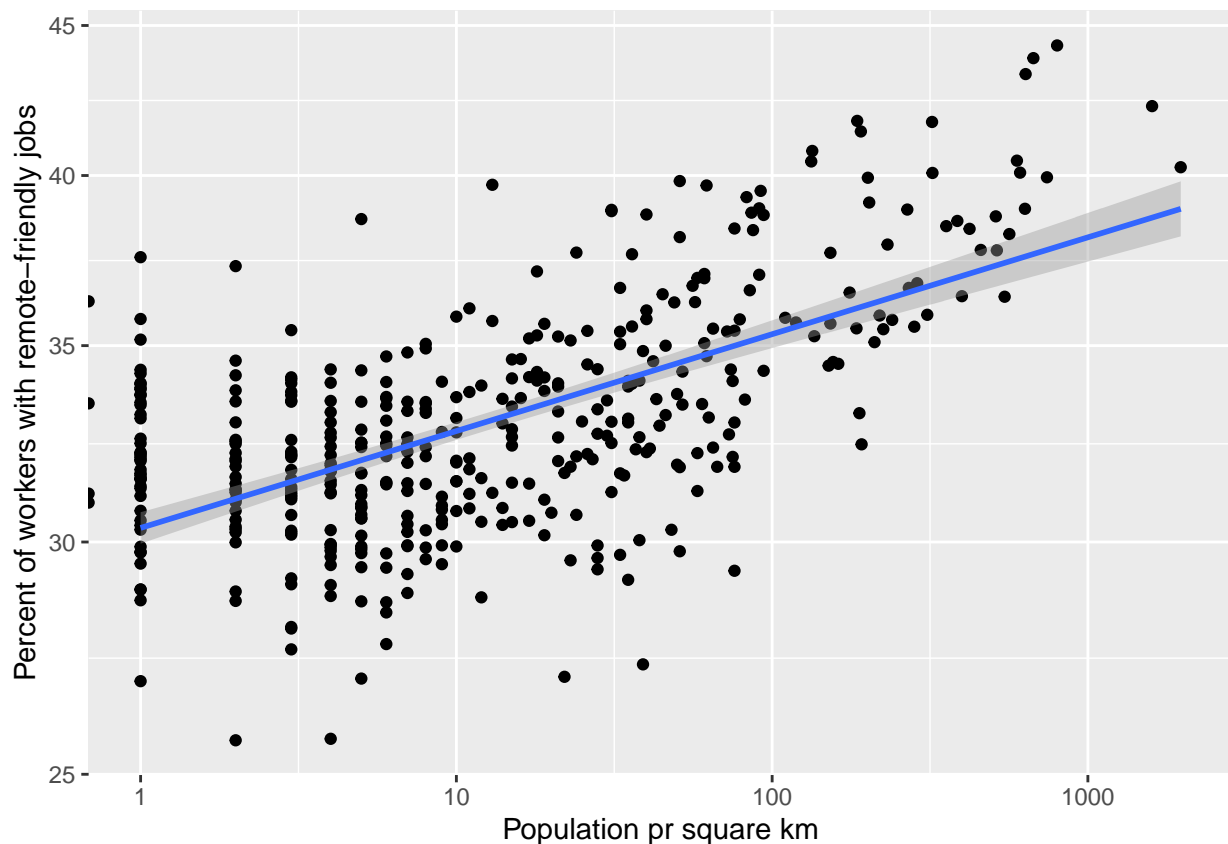
By introducing a measure of urbanness, we can analyze the relationship more formally. We use population per square km as a proxy for urbanness. From figure [TODO]figure:pop_density, we can see a clear correlation between urbanness, or population pr km^2, and the availability of remote-friendly jobs. Denser locations imply greater risks of Covid-19 spread, but this increased

risk may be mitigated by better opportunities for remote work.

European results

Employment by first-level ISCO codes are published by the european statistical agency Eurostat, and with a few assumptions these data can be used to estimate remote possibilities in the EU.

Due to the lack of 4-digit ISCO codes from Eurostat, we use the aggregates from Norwegian data as an estimate for remote percentages for 1-digit groups.

The countries with the highest share of remote-eligible jobs are rich countries, with Switzerland and Luxembourg topping the list.

```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Number of jobs (thousand)} & \\
##  \cmidrule(lr){2-3}
## Country & Total & Remote-friendly & Percent \\
## \midrule
## Switzerland & $4 477$ & $1 761$ & $39.34\%$ \\
## Luxembourg & $267$ & $103$ & $38.82\%$ \\
## United Kingdom & $30 943$ & $11 880$ & $38.39\%$ \\
## Norway & $2 565$ & $967$ & $37.68\%$ \\
## Sweden & $4 888$ & $1 830$ & $37.44\%$ \\
## Malta & $233$ & $86$ & $37.13\%$ \\
## \bottomrule
## \end{longtable}
```

On the other side, the countries with the lowest share of remote-eligible jobs are mostly poorer countries in southeast europe.

```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Number of jobs (thousand)} & \\
##  \cmidrule(lr){2-3}
## Country & Total & Remote-friendly & Percent \\
## \midrule
## Romania & $8 312$ & $2 077$ & $24.99\%$ \\
## Turkey & $27 882$ & $7 238$ & $25.96\%$ \\
## North Macedonia & $742$ & $197$ & $26.54\%$ \\
## Serbia & $2 671$ & $765$ & $28.64\%$ \\
## Bulgaria & $3 046$ & $883$ & $28.98\%$ \\
## Hungary & $4 398$ & $1 290$ & $29.34\%$ \\
## \bottomrule
## \end{longtable}
```
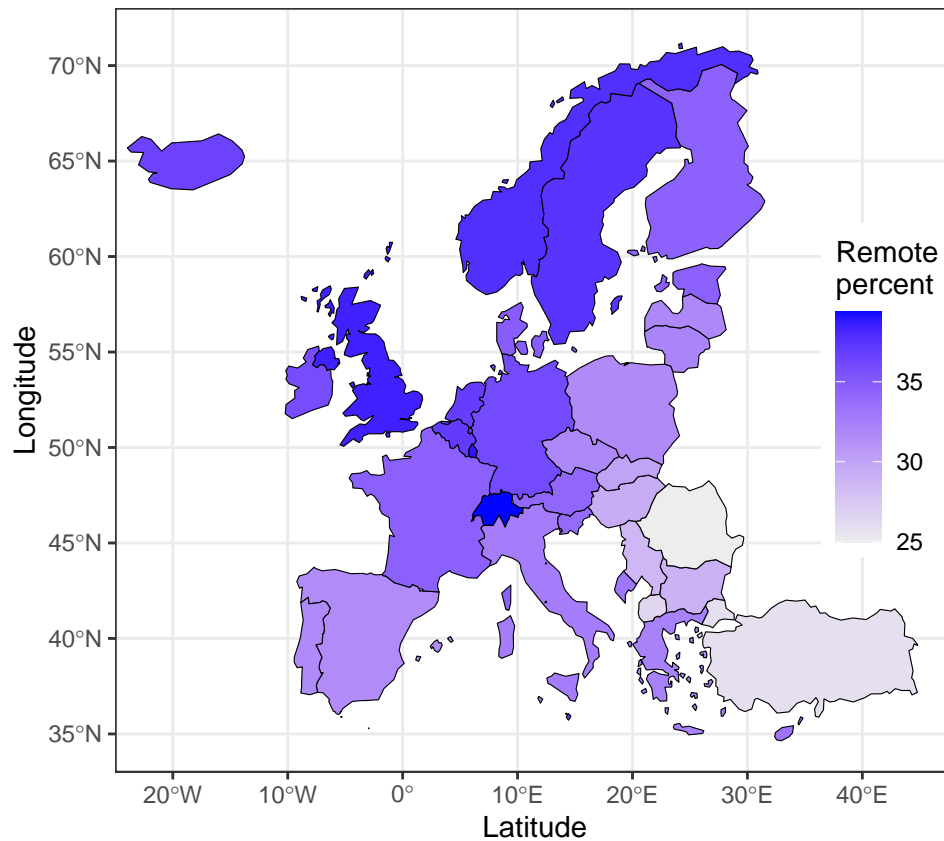
Validating results against job-ads

The job ads can not be compared directly to the annotated ISCO-08 data, but one of the comparisons we are able to do is the relative frequency of remote possibilities across occupational groups. Some job ads mention remote work, and we expect that remote possibilities are mentioned more often for occupations where remote possibilities are an option. We can compare this frequency with the number of remote-friendly occupations within each broader group (table tab:relative[TODO]). We also compute the composition of remote-friendly jobs across the occupational groups, giving us a measure that is directly comparable.

```
## \captionsetup[table]{labelformat=empty,skip=1pt}
## \begin{longtable}{lrrr}
## \toprule
## & \multicolumn{2}{c}{Relative remote frequency} & \\
##   \cmidrule(lr){2-3}
## Occupational group & Annotations & Job ads & Difference \\
## \midrule
## Managers & $3,9\%$ & $3,1\%$ & $0,8\%$ \\
## Academics & $67,6\%$ & $48,0\%$ & $19,6\%$ \\
## Assc. professionals & $13,3\%$ & $27,2\%$ & $-13,9\%$ \\
## Clerics & $3,5\%$ & $9,9\%$ & $-6,4\%$ \\
## Service workers & $5,8\%$ & $9,1\%$ & $-3,3\%$ \\
## Agricultural workers & $0,0\%$ & $0,0\%$ & $0,0\%$ \\
```

```
## Craft workers & $5,2\%$ & $1,2\%$ & $4,0\%$ \\
## NA & $0,5\%$ & $1,4\%$ & $-1,0\%$ \\
## NA & $0,2\%$ & $0,1\%$ & $0,2\%$ \\
## \bottomrule
## \end{longtable}
```

The notable discrepancies for the groups `Academics` and `Technicians and associate professionals` have a curious symmetry. Both of these groups require higher education or similar skill level, but Technicians and associate professionals are more vocational. There are many possible explanations for such differences, but for now it will suffice to conclude that one or more of the assumptions made is violated to some extent. Still the correlation is decent considering the spuriousness of the data.

In the appendix, we include some more discussions on this issue. It contains also a method we have developed to utilize job announcement data to correct for possible classification bias in the annotation labels. The magnitudes of the estimated biases are relatively small. Detailed estimates can also be found in appendix.

Related work

There are several very recent analyses that study the remote feasibility of jobs: (**?**), (**?**) and (**?**) for the United States, (**?**) for Germany, and (**?**) for Italy. Unlike our study, they mostly rely on different types of survey data. However, the results can be compared given the similar labor market institutions and labor force skill levels.

In Norway, as mentioned in section sec:introduction[TODO], there are two surveys related to our study. The results from the labor force survey was broadly similar to what we obtained with some differences. In the labor force survey, 71 per cent of managers responded that they had opportunity to work from home at times, much higher than the results here. This is likely attributable to the distinction mentioned above: Being able to perform *some* part of your job remotely does not mean that the job can be performed *primarily* remotely. The second survey by TØI was carried out after the Covid-19 outbreak in Norway. It reports an higher percentage of jobs (nearly 50%) are done from home these days. However, it is also reported considerable efficiency loss connoted to the home office arrangement. It may be that workers are stretching the limits of what can be effectively accomplished from home, or perhaps the annotations are slightly conservative. Note that although annotations were collected after the Covid-19 outbreak, it is hard to know whether the annotators had the pandemic in mind when annotating.

Conclusion

The sudden question of remote work highlights the need to expand our knowledge of occupations and their contents. In this paper we study the remote feasibility for different occupations in Norway. This analysis sheds more lights on a fundamental problem in the labor market. And more importantly it also provides useful knowledge for decision makers to evaluate potential social policies to combat the Covid-19 pandemic. Norwegian government will re-open primary schools and kindergartens soon. However, opening a school where majority of the parents can work from home would have rather different implications on potential virus spread, if compared with opening a school where majority of parents don't have this option. In addition, occupations

with different remote feasibility will not be hit as hard, so these estimates would be useful when one want to assess potential economic impacts of the pandemic.

On the other hand, for National Statistical Institutions in Europe, the most natural option would be to expand the existing ISCO ontology with this data. This analysis is an attempt to combine the conventional and unconventional sources for statisical and research purpose. The results we have found also suggest that alternative approaches to collecting such information is feasible. The ISCO-08 documentation includes more detailed descriptions of the jobs than what we made use of in this paper. The complete list includes lists of tasks commonly performed, which may further nuance what parts of a job can actually be performed from home. The ISCO ontology may also serve this purpose.Although there were some discrepancies between the job ads and the annotations, there is clearly a pattern, and the differences between such sources should be explored. From an economic perspective this may tell us something about employer preferences related to hiring, and from the perspective of a National Statistical Institution such sources may, if calibrated correctly, provide further information about the labor market and working conditions. The possibilities are not limited to the question of working from home. Job ads are a great resource for describing actual jobs and what they entail, although the picture the advertisements paint might be a little rosy.