Has COVID-19 changed task requirements on the job? Evidence from online job vacancies *

Ronald Bachmann[1,2], Niklas Benner[1], Rebecca Kamb[1], and Eduard Storm[1]

[1]RWI - Leibniz Institute for Economic Research

[2]IZA, Heinrich-Heine Universität Düsseldorf

February 2023

Abstract

This paper uses job vacancy data to provide causal analysis on the labor market effects of COVID-19. Using online job postings from January 2017 until December 2021 for the near-universe of German firms, we explore whether COVID-19 has changed task demand across firms and local labor markets. To permit causal interpretation, we exploit the sudden economic shock in March 2020 as a quasi-experiment and implement an event-study design coupled with Difference-in-Difference estimation. We find a short-term decrease in demand for interactive tasks by up to 12% and simultaneous increase in demand for manual tasks by up to 20% across local labor markets, with pronounced effects in agglomeration and rural regions. Our firm-level analysis adds important heterogeneities, suggesting these COVID-induced task shifts are primarily driven by (i) large firms, (ii) business related services, and (iii) firms with less technological adaptability to WFH.

Keywords: Job Vacancies, Event Study, COVID-19, Job Tasks, Future of Work

JEL Codes: D22, J23, J24, J63, O33, R11

*We are grateful to Eduard Brüll, João Pereira dos Santos, Lukas Hoernig and Sandra Schaffner for sharing code, data, and valuable insights. We thank all participants at the 3rd IAB/LISER Conference on “Labour markets during and after the COVID-19 pandemic” and the ZEW Workshop on “Big Data Analyses and New Developments in Research Data Centres” for valuable comments. Last, but certainly not least, we thank Anastasia Suckau for excellent research assistance. All remaining errors are our own.

1 Introduction

Technological change has caused large disruptions in labor markets in recent decades, causing job and wage polarization in many countries. This polarization has not only affected individual biographies (Blien, Dauth & Roth 2021), but also widened the gap between structurally advanced and weaker regions (Autor, Dorn & Hanson 2015). A consequence of these developments is changing task requirements on the job as firms are demanding new types of skills (Atalay, Phongthiengtham, Sotelo & Tannenbaum 2018, Bachmann, Cim & Green 2019). The analysis of online job vacancy (OJV) data using machine learning methods has contributed to our understanding about these mechanisms, especially by allowing researchers to exploit granular variation and performing near real-time analysis.

An influential series of studies has demonstrated the value of these methods by offering valuable ad-hoc insights about the impact of COVID-19 on labor demand. This literature has documented a sudden collapse and sharp subsequent rebound in job postings.[1] Many papers have built upon these insights to study various adjustment mechanisms. Some of these mechanisms are universal, e.g., the shift to remote work and subsequent questions about the "Future of Work".[2] Others are more country-specific, e.g., the important role of short-time work (STW) as coping mechanism in Germany.[3] Overall, these mechanisms are well-understood by now and point to important spatial differences depending on regions' Work-from-Home (WFH) capacity (Alipour, Falck & Schüller 2020) and digital infrastructure (Ben Yahmed, Berlingieri & Brüll 2022). Less-understood, however, is whether the pandemic-induced shift towards more digitalized workplaces has fundamentally changed the task content of jobs. Put differently, has the pandemic changed *what* type of tasks we perform or merely *how* we perform tasks (e.g., merely moving from performing tasks in-person to performing the same tasks virtually)? In this paper we fill this gap by exploring whether COVID-19 has caused shifts in task demand at the local labor market (LLM) and firm-level.

To answer this question, we use monthly OJV data from German firms to conduct a causal analysis based on an event-study design. This data has been scraped by our partner, a private IT company, covering the near-universe of job postings and spanning January 2017 until December 2021.[4] A novel aspect of this data is that we have access to the original

[1] See, e.g., Bamieh & Ziegler (2020), Campello, Kankanhalli & Muthukrishnan (2020), Forsythe, Kahn, Lange & Wiczer (2020*a*,*b*), Hensvik, Le Barbanchon & Rathelot (2021).

[2] See, e.g., Alipour, Falck & Schüller (2020), Alipour, Langer & O'Kane (2021), Dingel & Neiman (2020), Barrero, Bloom & Davis (2020, 2021), Adams-Prassl, Boneva, Golin & Rauh (2022).

[3] See, e.g., Adams-Prassl, Boneva, Golin & Rauh (2020), Ben Yahmed, Berlingieri & Brüll (2022).

[4] We are continuously receiving new batches of data. To date, we have data until June 2022. We have not included the recent batch in our papar yet as we are still cleaning and preprocessing the new data. We will, however, use it in our upcoming draft.

text data, allowing us to develop our own transparent taxonomy for task data and having complete control over the data-generating process. Our main finding is that demand for interactive tasks decreased by up to 6 pp. in the year 2020, a drop of 12% relative to the mean. In contrast, demand for manual tasks increased by up to 4 pp. in the year 2020, a rise of 20% relative to the mean. We attribute our key findings to the unique features of the pandemic. Lockdowns and subsequent contact restrictions prohibited many social activities and raised demand for "essential jobs", e.g., in healthcare or transportation/ packaging industries (Arthur 2021, Blau, Koebe & Meyerhofer 2021). Consequently, we find a reallocation of task demand away from interactive towards manual tasks, which are common in many industries that are deemed essential.

Analyzing task demand at the local level is important as aggregated data masks important heterogeneities, illustrated in Figure 1. Panel 1a displays aggregate task demand for five task groups and shows virtually no changes between 2017 and 2021. In contrast, Panel 1b displays the difference in task demand between regions at the 90th and 10th percentile. Prior to the pandemic, task demand differentials have converged, suggesting increasing similarity in the task content of jobs across regions. Since the pandemic started, however, task demand differentials have diverged, a development that has been driven by rising differences in demand for analytic and interactive tasks —activities that are common in many high-paying jobs, thus possibly exacerbating regional inequality.

We supplement our LLM-level analysis with firm-level evidence, allowing a richer characterization of changes in task demand. This micro evidence displays more pronounced heterogeneity at the firm-level, leading to comparably weaker responses in task demand relative to the LLM-analysis. To better understand underlying mechanisms, we study changes in task demand subject to firm characteristics. This analysis reveals that the reallocation away from interactive towards manual tasks has been primarily driven by (i) large firms, (ii) business related services, and (iii) firms with less technological adaptability to WFH.

Importantly, we show that a firm-level comparison between sectors masks the importance of firm-specific technology. Instead, technological adaptability to WFH only affects task demand *within* industries. We highlight this important distinction for firms in business related services, which includes professional services, IT services, and similar. Firms, which demonstrated their technological adaptability to WFH prior to 2020, show signs of upskilling. While they still experience a sizeable decrease in demand for interactive tasks by up to 5 pp., they also increase demand for analytic tasks by up to 2.5 pp. We do not find comparable evidence for any other type of firm, suggesting these "WFH firms" within business related services have been the most resilient firms in response to COVID.

[Figure 1 here]

Our paper contributes to several strands of the literature. First, it contributes to the literature that has quickly emerged at the beginning of the pandemic to document ad-hoc changes in labor demand.[5] A related strand of the literature studies the "Future of Work", often in the context of the shift to remote work and exploring how and whether WFH will transform workplaces.[6] We contribute to these papers by analyzing changes in task demand and exploring whether the pandemic has fundamentally changed the task content of jobs. Importantly, most of these papers provide descriptive evidence, whereas we are one of the few papers documenting causal evidence on the impact of COVID-19.

Second, our paper is also closely related to a few papers studying regional consequences of the pandemic using causal methods. For example, Carvalho, Peralta & Pereira Dos Santos (2021) and Hoseini & Valizadeh (2021) study changes in purchasing behavior in Portugal and Iran, respectively, using transaction data. Studying a similar setting like ours, Ben Yahmed, Berlingieri & Brüll (2022) explore differential uptake of short-time work across German regions depending on their pre-pandemic digital infrastructure and differences in industry composition. We contribute to this literature the causal analysis of regional differences in labor demand using firm-level OJV data as well as firm-level evidence on the causal impact of the pandemic.

Third, and more broadly, our paper contributes to the large task literature. Many papers use survey-based task data to study worker-level differences in wages or employment.[7] Even more papers study changes in task demand between and within occupations over time.[8] These studies suffer from small sample limitations or lagged information as timely task data is usually not available. Our overarching contribution to this literature is the use of "Big Data" as we have 22 million observations on task demand, allowing us to overcome small sample limitations (especially at the spatial dimension), and near real-time data until June 2022, allowing us to perform a timely analysis of economic shocks. Moreover, we have the original text data, allowing us to have full control over the data-generating process. This is an important distinguishing feature from the only other paper we know using task information from OJV data —data that has been preprocessed and preclassified (Blanas & Oikonomou

[5] See, e.g., Bamieh & Ziegler (2020), Campello, Kankanhalli & Muthukrishnan (2020), Forsythe, Kahn, Lange & Wiczer (2020*a*,*b*), Marinescu, Skandalis & Zhao (2020), Shen & Taska (2020), Hensvik, Le Barbanchon & Rathelot (2021).

[6] Notable papers include Alipour, Falck & Schüller (2020), Alipour, Langer & O'Kane (2021), Dingel & Neiman (2020), Barrero, Bloom & Davis (2020, 2021), Adams-Prassl, Boneva, Golin & Rauh (2022), de Fraja, Matheson & Rockey (2021).

[7] See, e.g., Spitz-Oener (2006), Autor & Handel (2013), Cassidy (2017), Rohrbach-Schmidt (2019), Storm (2022*a*,*b*).

[8] Notable papers include Gathmann & Schönberg (2010), Atalay, Phongthiengtham, Sotelo & Tannenbaum (2018), Bachmann, Cim & Green (2019), Bachmann, Demir & Frings (2021), Cortes, Lerche, Schoenberg & Tschopp (2021), Cortes, Jaimovich & Siu (2021), Bachmann, Demir, Green & Uhlrendorff (2022).

2022).

Fourth, and lastly, we contribute to the literature on recessions and recruitment. Deming & Kahn (2018) show variation in regional skill requirements is informative about differences across local labor markets. We complement their study by studying variation in regional task demand. Importantly, several papers show the Great recession 2007/08 has accelerated polarization as regions with higher unemployment rates increase skill requirements (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019). This "upskilling" is attributed to a larger pool of skilled applicants from which firms can choose to perform cognitively demanding tasks and has primarily hurt young workers (Forsythe 2022).

We contribute to this literature by studying changing task requirements in response to COVID-19. We think this distinction is important as experiences from the Great Recession may not readily apply to the current pandemic due to its unique features (e.g., lockdowns, social restrictions). Most closely related to our paper is an important recent study by Blanas & Oikonomou (2022). They offer descriptive evidence on the changes in demand for occupations due to pandemic-induced uncertainty in the US until December 2020. We contribute to their study by offering causal analysis on the impact of COVID-19 on task demand and providing firm-level evidence.

2 Data

2.1 Online Job Vacancies

2.1.1 Data-Generating Process

We use a unique data source consisting of the near-universe of online job vacancies posted in Germany between January 2017 and June 2022. The job postings are collected by our private partner, a firm that is offering custom-made firm-, person- and job posting- data and market analysis. Our partner scrapes more than 2,000 web-pages for vacancies from the following platforms: (i) job boards (fee paying), (ii) job boards (free of charge), (iii) company websites, (iv) temporary employment agencies, and (v) head-hunters. They consistently update their online sources and scrape all sources on a daily basis. Subsequently, our partner performs some basic cleaning procedures, such as removal of "boilerplates" (i.e., content that is unrelated to the vacancy, such as ad text) and removes duplicates from the same source (i.e. sources from the same url address). A unique feature of this data is that our partner merges posting firms with the German company registry ("Handelsregister"). This merge is successful for about 60% of firms.[9] Subsequently, our partner sends us the data, outlined below.

One, we receive firm information, containing the location of the firm (in many cases at zip-code level), industry, and other important background information. Two, we receive person information, such as sex, function, and date of birth of the owner. Three, we receive source information, i.e. the type of platform the vacancy was scraped from. Four, we receive vacancy information, containing the original job description.

Upon receiving the data, we perform further cleaning procedures. First, we link firm and vacancy information, especially in order to assign job descriptions to specific industries. Second, we use this linked data set to assign OJV to specific locations, preferably at the zip-code level. When zip-code level data is not available, we use information on the job site (i.e., city, town, or village).[10]. For about 10% of vacancies we only observe the job site at a broader level, e.g., district-level. For the purpose of this paper we omit these observations and only use zip-code and job site information as we seek sufficiently precise information about the location of the workplace. Third, we create a unique taxonomy of tasks, containing various single activities. We describe this taxonomy and its consistency with the existing literature in more detail in section 2.3.

For our analysis, we focus on vacancies for regular work, i.e. full- or part-time. Thus, we remove vacancies seeking apprenticeships, trainees, and other types of irregular work. In particular, we drop vacancies for temporary employment as they are likely not representative of regular labor market developments (Stops, Bächmann, Glassner, Janser, Matthes, Metzger & Müller, Christoph, Seitz, Joachim 2021). Temporary employment agencies are special in the sense that their postings may be counter-cyclical: If labor demand is small, they may increase the number of persons in their applicant pool, and publish less postings if labor demand is high in the labor market. Therefore, job vacancies of temporary employment agencies distort demand for labour demand and show patterns that are incompatible with official statistics. Moreover, we keep only those vacancies for which we have firm-level

[9]The data set is based on information from the Handelsregister and includes all firms that are listed in the Handelsregister since 1991. About half of the 3,4 Mio. firms in Germany are noncommercial and therefore not listed in the Handelsregister. In addition, firms from the public administration sector are not included. The firm level data includes information about the firm name, the complete address, legal status, industry, original stock and business volume, the number of employees and the formation date. The data can be merged through a firm identifier, which is available for about 60% of the job postings. Reasons why the firm identifier is not available are, on the one hand, that firms are not listed in the Handelsregister, or, on the other hand, because group of companies cannot be assigned to one specific firm.

[10]We know the zip code of the job site for about 60% of vacancies and the name of the city, town, or village for about 30% of vacancies. In some cases, two job sites have the same or vague names, e.g., *Frankfurt am Main* versus *Frankfurt an der Oder*. In these cases, we assign the vacancy to both cities and weigh both observations by half.

information from the company registry. This way, we maintain a consistent sample for our analysis at the LLM- and firm-level.

After cleaning and selecting the relevant data, we are left with 22 million job vacancies, comprising 240,000 firms and 2.1 million firm-month observations. In a final step, we perform a few more standard preprocessing steps on the job description. Specifically, we follow Gentzkow, Kelly & Taddy (2019) and preprocess the text data for the empirical analysis by (i) converting job descriptions with tokenization, (ii) removing stop words, and (iii) stemming words.

In section D.1 in Appendix D we provide external validity on our data quality by comparing trends by time and industries to official job vacancy statistics. Overall, we demonstrate our OJV data depicts similar trends between 2017 - 2021 and covers all industries properly.

2.2 Regional Data and Local Labor Market Definition

We supplement our OJV data with regional characteristics to account for systematic differences between LLMs that may confound our analysis on local changes in task demand. These variables are taken from a regional administrative database, Regionalstatistik.de (Regionalstatistik 2022/10/25), and comprise various statistics at the county-level, such as local skill composition, socio-economic composition (age, gender, citizenship), industry composition, and the average local unemployment rate.

We perform our analysis at a broader definition than county-level. Counties have administrative boundaries that do not necessarily reflect LLM in an economic context. For instance, counties do not account for common commuting zones. Disregarding these movements may introduce spillovers and thus bias our results. We therefore aggregate the 402 counties into 141 broader LLM, following the classification of Kosfeld & Werner (2012), which has been used widely in research on LLMs in Germany (Dauth, Findeisen, Suedekum & Woessner 2021, Dustmann, Lindner, Schönberg, Umkehrer & vom Berge 2021, Hirsch, Jahn, Manning & Oberfichtner 2022).

2.3 Task Data

Our access to the original texts of the vacancies allows us to have complete control over the data-generating process and develop our own, transparent task taxonomy. In contrast, the existing literature uses task information that has been preprocessed and classified by the respective data provider (Blanas & Oikonomou 2022). For our taxonomy we collect job activities that have been frequently adopted in the existing task literature, either based on survey responses[11] or retrieved from the online portal BERUFENET, the German equivalent of the US O*NET database[12]. As we use our keyword list to scan through job postings, we aim to have a sufficiently comprehensive list of activities. For this reason, we supplement these activities with equivalent descriptions of the same task in order to expand the scope of our search.

Subsequently, we follow the literature (Autor, Levy & Murnane 2003, Spitz-Oener 2006, Storm 2022*b*) and classify a variety of single activities into five broad task categories: (i) non-routine (NR) analytic, (ii) NR interactive, (iii) routine (R) cognitive, (iv) R manual, and (v) NR manual. NR analytic and NR interactive involve strong problem-solving skills and abstract thinking. In contrast, routine tasks are characterized by following explicit and easily codifiable rules. Lastly, NR manual requires physical labor pronounced in, for example, basic services.

NR analytic and NR interactive involve strong problem-solving skills and abstract thinking. In contrast, routine tasks are characterized by following explicit and easily codifiable rules. Lastly, NR manual requires physical labor pronounced in, for example, basic services.

Figure 2 displays word clouds, illustraing the most important activities belonging to each of the above task groups. Overall, we consider the most important activities within each task group intuitive and in line with existing job task descriptions used in the literature. For instance, *analysis* and *development* are the two most important activities within the task group NR analytic, while activities in the realm of *human resource management* are the most important activities within the task group NR interactive. Regarding the routine task groups, *register* and *surveillance* are the most important activities within Routine cognitive and *preparation* and *production* are the most important activities within Routine manual. Lastly, *equipping* [machines] and *support* are the most important activities within NR manual.

[Figure 2 here]

3 Methodology

3.1 Task Construction

In this section, we describe the construction of our outcome variables, task intensities T_{ijlmt} and T_{jlmt}, using our task taxonomy described above. First, we count for each firm i located in LLM l the number of online vacancies that have been posted in month m in year t. Second,

[11] For an overview of tasks used in the survey-based literature, see, e.g., Spitz-Oener (2006), Gathmann & Schönberg (2010), Rohrbach-Schmidt & Tiemann (2013), Storm (2022*b*,*a*).

[12] For an overview of tasks used in this literature, see, e.g., Dengler, Matthes & Paulus (2014).

we count the number of times that task j has been posted across all vacancies. Third, we compute the average task intensity for each firm i by dividing the number of tasks by the number of overall vacancies posted.

For example, if firm i in LLM l posted 10 vacancies in January 2017 and we count a total of 50 NR interactive tasks in these postings, the absolute task intensity of NR interactive in i in January 2017 was 5. Performing this calculation for each of the five tasks gives us the full distribution of task intensities. To alleviate concerns regarding trends in firms' posting behavior - e.g., more postings, see Figure 19, or more task requirements - we calculate the relative task intensity T_{ijlmt} as follows:

$$T_{ijlmt} = \frac{\text{Number of tasks j demanded by firm i in LLM l in month m in year t}}{\text{Total number of tasks demanded by firm i in LLM l in month m in year t}} \qquad (1)$$

where $j = 1, ..., 5$ represents the five tasks defined above. This definition implies (i) $T_{ijlmt} \in [0, 1]\ \forall j$ and (ii) $\sum_J T_{ijlmt} = 1$, thus describing the average relative importance of each task j in vacancies posted by firm i. For example, $T_{NRI,ilmt} = 0.5$ implies 50% of all tasks demanded in vacancies by firm i at time t are interactive. Variations of this measure have been widely adopted in the task literature, making our results comparable to previous research (Antonczyk, Fitzenberger & Leuschner 2009, Bachmann, Demir & Frings 2021, Storm 2022*b*,*a*). For our LLM-analysis we simply calculate the average task intensity across all LLMs, i.e. $T_{jlmt} = \frac{1}{l} \sum T_{ijlmt}$.

3.2 Identification

In our baseline identification strategy we follow Carvalho, Peralta & Pereira Dos Santos (2021) and Hoseini & Valizadeh (2021) by assuming a binary treatment using the onset of the pandemic in March 2020 as a quasi-experiment. Following their methodology, our baseline analysis also takes place at the LLM-level, i.e. our baseline outcome variable is T_{jlmt}. We consider the year 2020 as the treatment period and each of the months from 2020/03 - 2020/12 as treatment units. The months 2020/01 - 2020/02 serve as our control units. Our identifying assumption is that the year-on-year (YoY) change in T_{jlmt} between each of the treatment units, 2020/03 - 2020/12, and the preceding months, i.e., 2019/03 - 2019/12, would be parallel to the YoY change between January/ February 2020 and January/ February 2019 —had the pandemic not occurred.

We check the validity of this this assumption by studying the pre-COVID periods in our event-study specification discussed below. In further robustness checks we extend the pretrend period period and perform various other tests. A caveat of this identification strategy is that it allows to draw causal inference only for the year 2020 and thus merely a short-term analysis. The reason is that an extension of the parallel trends assumption would be violated otherwise as the YoY change of months in 2021 and preceding months in 2020 would include months from 2020 that had been treated. We aim to address this shortcoming with an alternative identification strategy in our next draft.

3.3 Empirical Methodology

We perform a standard Event-Study Difference-in-Difference (DiD) estimation. Our LLM-analysis closely follows Carvalho, Peralta & Pereira Dos Santos (2021) and is described as follows:

$$\begin{aligned}\underbrace{T_{jlmt}}_{\text{Task Demand}} &= \underbrace{\lambda_l + \mu_m}_{\text{Two-Way FE}} + \underbrace{\alpha \times 1_{t=2020}}_{\text{Treatment Period}} \\ &\underbrace{\sum_{m=-2}^{M=10} \beta_m \times 1_{t=2020} \times 1_m}_{\text{Treatment Units (m} \geq 0\text{), Control Units (m} < 0\text{)}} + \delta \mathbf{X}_{lmt} + \epsilon_{jlmt}\end{aligned} \tag{2}$$

where j $\in (1, ..., 5)$ reflects task groups, l $\in (1, ..., 141)$ reflects labor market regions, m $\in (1, ..., 12)$ reflects calendar months, and t $\in (2017, ..., 2020)$ reflects years. λ_l denotes LLM fixed effects and μ_m denotes monthly fixed effects to account for regional and time trends. The vector $\mathbf{X}_{lmt}$ comprises various controls at the LLM-level to account for confounding factors that may affect local task demand.[13] We are interested in coefficients β_m $\forall m \geq 0$ as they capture the treatment effects of the impact of COVID-19 on local task demand T_{jlmt} for each month from March 2020 - December 2020. Carvalho, Peralta & Pereira Dos Santos (2021) show formally that, given our identifying assumption and empirical setup, β_m is an estimate of the difference in local task demand between 2019 and 2020 for each m and the a weighted geometric average of the YoY growth rates of the preceding two years. The constant is omitted from equation (2) for brevity. Moreover, we define February 2020 as reference unit. All coefficients must thus be interpreted relative to the month before the pandemic started. We cluster standard errors at the state-level to account for serial correlation at more aggregated levels.

[13] We include the following controls: share of college graduates, share of workers with completed vocational schooling, share of workers with neither ob above schooling requirements, share of different age groups (six age bins), share of workers with foreign citizenship, share of female workers, industry composition (1-digit, 13 industries), unemployment rate, technology differences (measured via share of vacancies in LLM that offer WFH option).

For our firm-level analysis, we estimate a modified version of (2), using T_{ijlmt} as outcome variables. This analyis is based on the following specification:

$$\underbrace{T_{ijlmt}}_{\text{Task Demand}} = \underbrace{\lambda_i + \mu_m}_{\text{Two-Way FE}} + \underbrace{\alpha \times 1_{t=2020}}_{\text{Treatment Period}} \underbrace{\sum_{m=-2}^{M=10} \beta_m \times 1_{t=2020} \times 1_m}_{\text{Treatment Units (m} \geq 0\text{), Control Units (m} < 0\text{)}} + \delta \mathbf{X}_{ilmt} + \epsilon_{ijlmt} \tag{3}$$

In contrast to our LLM-specification, our firm-level model incorporates firm-level FE and clusters standard errors at the LLM-level. Moreover, we include firm-level covariates, including firm age, employment, and revenue. Otherwise, both specifications are alike.

4 Baseline Results

4.1 LLM-Analysis

We begin our discussion of results with Figure 3, illustrating the baseline event studies for the demand of each our tasks at the LLM-level. Each panel highlights the first wave of the pandemic starting in March 2020 ("COVID-Lockdown") and the second wave that has caused further lockdown restrictions in November 2020 ("Lockdown-Light"). As a first step, we inspect pre-trends to gauge the validity of the parallel trend assumption. For all our specifications we cannot reject the hypothesis that the pre-pandemic coefficients $\beta_m \; \forall m < 0$ are different from zero. This observation lends credence to our claim of causal effects. Of course, however, we cannot rule out confounding factors or misspecification that are masked in our baseline model. We address these concerns in our robustness checks in section 4.3.

Panels 3a and 3b display the baseline results for NR analytic and NR interactive tasks, both of which are cognitively demanding. Demand for NR analytic tasks has remained steady throughout the year 2020, displaying no COVID-induced change in demand for analytic tasks. In contrast, we observe a strong and steady decrease in demand for NR interactive tasks from June 2020 onward. This decline reached its peak in fall 2020, when infection rates started to climb again during the second wave, with a reduction in demand of interactive tasks of up to 6 pp. Relative to the mean of 50% in NR Interactive task intensity (Figure 1), this drop corresponds to a sizeable 12% decline in demand for NR interactive tasks. Considering these tasks are highly social and most important in jobs in pedagogy and retail, it is not surprising their demand was heavily affected by contact restrictions that were common throughout the year 2020.

[Figures 3a - 3e here]

Panel 3c displays the baseline results for Routine cognitive tasks, which are common in many office and clerical jobs. Demand for these tasks has remained flat for most of the year 2020, except for a short spike during summer 2020. Panels 3d and 3e display the baseline results for Routine and NR manual tasks. Demand for both of these tasks has increased throughout the year 2020, however, in different cycles. Demand for routine manual tasks has been more consistent and sizeable, reaching an increase of up to 3 pp. during "Lockdown-Light" in Fall 2020 and reflecting an increase of 30% relative to the mean. We attribute part of this substantial increase to the rising importance of online sales as retail stores were forced to shut down initially and were allowed only a limited amount of customers for other parts of the year (Goecke & Rusche 2022). This hypothesis is intuitive considering routine manual tasks are most common in the transportation/ packaging and manufacturing industries and is further reflected in the importance of single activities such as preparation, packaging, and shipping within the routine manual task group (Figure 2d).

In contrast, demand for NR manual tasks was more episodic throughout the year. At the onset of the pandemic in March 2020, demand for NR manual task increased immediately by up to 1 pp., or, 10% relative to the man. Considering these tasks are most pronounced in the health industry, these findings are unsurprising, partially reflecting rising demand for nursing staff. As infection rates decreased in summer 2020, demand for NR manual tasks likewise decreased. While demand was somewhat volatile in the months to come, it picked up again shortly in fall 2020 when the second wave started, likely reflecting demand for "frontline workers" (Blau, Koebe & Meyerhofer 2021) as these tasks are pronounced in essential jobs such as caretaking and health services. Consistent with this notion, Arthur (2021) finds a COVID-induced increase in vacancies associated with care work and nursing jobs in the UK. Likewise, Blanas & Oikonomou (2022) document an increase in demand for contact-intensive essential jobs in the US, especially those jobs that are service-related.

We find our baseline results credible and intuitive as they reflect features that were unique to the pandemic. On the one hand, lockdowns and contact restrictions are most likely to detrimentally affect jobs in service industries that require many personal interactions. On the other hand, the pandemic has been characterized by episodic changes in infection rates, which in turn affected demand for certain essential jobs and thus demand for tasks. However, our results are partly at odds with evidence from the US. Blanas & Oikonomou (2022) find an increase in demand for occupations intensive in NR analytic and NR interactive tasks and a decrease in demand for occupations intensive in customer relationships. Yet, unlike our study, Blanas & Oikonomou (2022) study state-level variation, possibly masking underlying heterogeneity at the LLM-level.[14]

Indeed, when we break down our LLM-analysis by large agglomeration areas, mid-sized urban areas, and rural areas, we detect important heterogeneites.[15] Notably, the decrease in demand for NRI tasks has been primarily driven agglomeration and rural areas. Intuitively, agglomeration areas have a higher population density, making contact restrictions more binding and thus interactive tasks more difficult to facilitate. In contrast, rural areas tend to be less digitalized and thus are not able to shift work remotely. This lack in digital capital may prevent a stabilizing force in the decline in demand for social tasks. To validate these hypotheses and add micro-foundations to these LLM-level insights, we now move to our firm-level analysis.

[Figures 4 - 6 here]

4.2 Firm Heterogeneity

In this section, we inspect the regional heterogeneities in task demand in more detail by exploiting firm-level variation. To this end, we (i) repeat our baseline analysis at the firm-level, and moreover study changes in task demand subject to (ii) firm size, (iii) industry affiliation, and (iv) technology, measured by the prevalence of WFH options pre-pandemic.

A. Baseline

Our findings at the firm-level are summarized in Figure 7, which are based on model (3). We first inspect pre-trends to gauge the validity of the parallel trend assumption. Again, we cannot reject the hypothesis that the pre-pandemic coefficients β_m $\forall m < 0$ are different from zero, lending credence to our causal interpretation.

Qualitatively, firm-level results are not substantially different compared to our previous LLM-level analysis. The key difference is in regards to demand for NR manual and interactive tasks. Recall that our LLM-analysis shows an immediate increase in demand for NR manual tasks at the onset of the pandemic. In contrast, we find no changes in demand for NR manual tasks at the firm-level until August 2020. Hence, the micro evidence in this section

[14]Moreover, we believe part of this discrepancy can be attributed to different strategies in combatting COVID-19. While the US allowed a high unemployment rate of up to 15%, instead providing generous income stimulus, Germany relied heavily on a short-time work model ("Kurzarbeit"), thereby capping unemployment at 6%. Based on the experiences from the Great Recession 2007/08, upskilling (in form of higher analytic and interactive task requirements) is more likely to occur in regions with high unemployment rates (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019). Therefore, COVID-induced regional variation in unemployment was *less* pronounced in Germany, suppressing upskilling incentives.

[15]We follow Hamann, Niebuhr, Roth & Sieglen (2022) by dividing LLMs into these three distinct groups and thank the authors for making their classification available to us.

suggests more pronounced firm heterogeneity at the beginning of the pandemic, during times of heightened uncertainty.

Quantitatively, our firm-level analysis moreover indicates a weaker decrease in demand for interactive tasks by up to 5 pp. (compared to 6 pp. at the LLM-level). In the following, we inspect these discrepancies more closely by studying changes in task demand subject to firm characteristics.

[Figures 7a - 7e here]

B. Firm size

In this exercise, we split our sample of firms into three groups: (i) large, (ii) mid-sized, and (iii) small firms. We define firms as "large" if their workforce is in the top quartile of the firm size distribution ($>$ 275 employees). Similarly, we define firms as "small" if their workforce is in the bottom quartile of the firm size distribution ($<$ 16 employees). The remaining "mid-sized" firms are located in the interquartile range of the firm size distribution (16 - 275 employees). The results are reported in Figures 8 - 10.

Our key takeaway from this exercise is that the baseline findings are largely driven by large firms (Figure 8). In comparison, smaller firms had modest declines regarding demand for interactive tasks relative to large firms. Specifically, Mid-sized firms with 16 - 275 employees experienced a peak decline in demand for NR interactive tasks by up 2.5 pp. in fall 2020. In comparison, small firms with less than 16 employees faced a peak decline by up to 5 pp. in July 2020. Notably, small firm's demand for NR interactive tasks converged back to pre-COVID levels in subsequent months - a pattern we do not observe for larger firms.

Combined, our analysis suggests that larger firms changed their task requirements more strongly away from interactive towards manual tasks. Considering large firms are more susceptible to contact restrictions by virtue of employing more people, we suspect lockdown measures had stronger bite for these firms. Consequently, regions containing a disproportionate number of large firms experienced sharper changes in task demand, consistent with the regional heterogeneities discussed in section 4.1. However, given the nature of the pandemic, we expect sizeable differences across industries, which we explore in the next section.

[Figures 8 - 10 here]

C. Industry affiliation

We repeat our baseline analysis separately for 8 broad industries: (i) agriculture and mining, (ii) food and consumer goods, (iii) industrial goods, (iv) capital goods, repair, and installation of machines, (v) construction and provisioning, (vi) personal services, (vii) business related services, and (viii) public sector.

For brevity, we will focus on personal and business related services in this section, although all other results are available upon request and will be made available in an upcoming online appendix. We choose these two industries to highlight heterogeneous responses across industries. Personal services comprise travel industry, event management industry, and other related industries that were effectively shut down at the beginning of the pandemic. Business related services, in turn, comprise consulting, marketing, and scientific R&D; i.e. (service) industries that rely on human interaction but were still operating throughout the pandemic.

Figures 11 - 12 summarize the results of this exercise. Somewhat surprisingly, we find no decrease in demand for interactive tasks in personal services throughout 2020. In comparison, business related services experienced a pronounced decline in demand for interactive tasks by up to 5 pp., thus in accordance with our (firm-level) baseline results. This drop in interactive tasks in business related services was compensated by increasing demand for NR manual tasks, by up to 3 pp., and routine manual tasks, by up to 2 pp.

[Figures 11 - 12 here]

At first glance, these results may seem counterintuitive. Considering business related services include more tasks that can be performed remotely than personal services, including interactive tasks (Alipour, Falck & Schüller 2020, Dingel & Neiman 2020, Adams-Prassl, Boneva, Golin & Rauh 2022), shouldn't this industry then demonstrate greater resilience regarding its task demand? We believe two reasons can explain these results.

First, task demand in personal services is relatively less elastic than task demand in business related services. We support this conjecture with pre-pandemic differences in task demand. Prior to the pandemic, 47% of all task requirements in personal services were of interactive nature. These tasks are highly sociable and cannot easily be performed without human interaction, e.g. hospitality. In comparison, only 40% of task requirements were interactive in business related services. Instead, firms in business related services put greater emphasis on analytic tasks (30% vs 23% in personal services). On average, task intensity with respect to interactive tasks even increased to 49% for firms within personal services, while remaining flat for firm within business related services. Hence, tasks in personal services are more dependent on human interaction and seem less responsive to exogenous shocks.

Second, we believe part of differential results between these industries is attributed to technological factors, which are masked in this exercise. In particular, COVID has induced higher levels of digitalization, most notably in form of greater prevalence of WFH. Firms being able to have their employees work remotely are likely exposed to more digital technologies, which in turn may have stronger implications on their task demand. Firms in business related services are more exposed to these technologies, in part by containing more IT firms. We test this hypothesis in the next section.

D. Technology (WFH)

In our last firm-level exercise, we explore the technological implications of the COVID-19 shock. The pandemic imposed greater reliance on WFH on many firms, thus inadvertently creating an experiment to test their adaptability to remote work. We approximate these technological differences by dividing the sample into firms that already offered WFH prior the pandemic and those firms that did not. Following the literature, we consider WFH to be a viable proxy for digital infrastructure and thus a proxy for technological feasibilty for WFH, similar to Ben Yahmed, Berlingieri & Brüll (2022). We then repeat our event study for both types of firms. In what follows, we refer to firms that offered WFH prior to 2020 as “WFH Firms” and those that did not as “non-WFH Firms”.

[Figures 13 - 14 here]

Figures 13 - 14 summarize the results of this exercise. We find no discernible differences on task demand between “WFH Firms” and “non-WFH Firms”. Both types of firms display changes in task demand consistent with baseline findings. Surprisingly, the decrease in demand for NR interactive tasks has been even more pronounced among “WFH Firms”, being about twice as large by the end of 2020 compared to “non-WFH Firms”. This exercise thus suggests that prior WFH experience has no effect on firm’s task demand. However, WFH offers were very limited prior to the pandemic, with only 10% of firms making such offers pre-2020. Even in business related services, an IT-affine industry, only 15% of firms offered WFH prior to 2020. Since above exercise compares firms *between* industries, overall limited prevalence of WFH pre-pandemic may mask within-industry heterogeneities.

For this reason, we supplement above analysis by differentiating between “WFH Firms” and “non-WFH Firms” *within* industries. Specifically, we focus on business related and personal services, as in the previous section.

[Figures 15 - 18 here]

Figures 15 - 18 summarize the results of this exercise. They key finding is on the adjustment mechanism of “WFH firms” in business related services (Figure 15). While these firms still display a sizeable decrease in demand for interactive tasks by up to 6 pp., they make up for it by greater emphasis on analytic tasks. At the beginning of the pandemic, “WFH firms" increased their demand for analytic tasks by 1 pp., or 5% relative to the mean. By fall, around Lockdown Light, demand even increased by up to 2.5 pp., or 12.5% relative to the mean. This observation is in sharp contrast to "non-WFH firms" within the same industry. These firms had comparable declines in demand for interactive tasks, yet increased their demand for routine and NR manual tasks to a greater extent instead. We do not observe similar adjustment mechanisms among firms within personal services (Figures 17-18), nor do we observe it among any of the other industries.

Combined, our evidence suggests technological adaptability affects task demand only *within* industries. According to this interpretation, "WFH firms" within business related services have been the most resilient firms in response to COVID. Less need for physical presence at the workplace combined with prior experience with WFH allowed these firms to adjust their task requirements more broadly. Next to more demand for manual tasks, these firms also used the COVID-imposed shift in digitalization to place greater emphasis on analytic tasks, i.e., a form of upskilling frequently observed in the literature on skill requirements following the Great Recession (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019).

4.3 Robustness

We perform a series of tests to check the validity of our event studies, specifically at the LLM-level. First, we change the definition of LLMs. While our baseline definition of LLMs is rather broad, comprising 141 regions, this definition may mask underlying heterogeneity.[16] We re-run our analysis, using a more narrow definition of LLMs instead, comprising 257 regions, that has also been used in the previous literature, see, e.g., Ben Yahmed, Berlingieri & Brüll (2022). Figure 21 shows our key results remain robust to this alternative definition of LLMs.

Our next robustness test addresses concerns regarding unobserved regional seasonality that may be driving our results. Figure 22 illustrates the results for an exercise in which two-way FE (month and LLM) are replaced by month $\times$ LLM fixed effects. Again, our results remain qualitatively unaffected by this robustness check. Similarly, changing clusters from the state-level to the LLM-level to test for alternative correlation of standard errors leaves our findings unaffected (Figure 23). Lastly, we test for misspecification of our event-study design by extending the pretrend period by two more months (Figure 24. This exercise also leaves our findings unaffected, lending credence to our baseline methodology.

[16] There are 402 counties in Germany. Therefore, our baseline definition of LLMs aggregates almost three counties to one LLMs, on average.

[Figures 21 - 24 here]

5 Conclusions

In this paper we use monthly online job vacancies data to study the causal impact of COVID-19 on changes in task demand across 141 German local labor markets (LLM) and 240,000 firms. We perform this analysis with a event study design coupled with Difference-in-Difference estimation, using the sudden pandemic-induced restrictions in March 2020 as a quasi-experiment. Our key finding is a sizeable short-term shift in task demand away from interactive towards manual tasks throughout the year 2020. At the LLM-level, demand for interactive tasks decreased by up to 6 pp. in 2020 —10% relative to the mean —reflecting the prohibitive implications of social restrictions for many jobs. At the same time, demand for manual tasks increased by up to 4 pp. in 2020 —20% relative to the mean —reflecting the increasing demand for "essential jobs" (Arthur 2021, Blau, Koebe & Meyerhofer 2021). Our evidence demonstrates an important reallocation in demand for tasks that is likely attributed to the unique features of the pandemic, namely lockdown measures and contact restrictions.

Our firm-level evidence displays weaker task shifts compared to the LLM-level, pointing to more pronounced heterogeneity within local labor markets. In subsequent analysis, we use this firm-level variation to shed more light on underlying mechanisms. This approach reveals that the reallocation from interactive towards manual tasks has been primarily driven by (i) large firms, (ii) business related services, and (iii) firms with less technological adaptability to WFH. An important insight from this analysis is firm heterogeneity within industries, which is especially pronounced in business related services. Firms that have demonstrated their technological adaptability to WFH prior to 2020, by virtue of pre-pandemic WFH offers, compensate half of the decline in interactive tasks by more demand for analytic tasks. We do not find such an observation for any other type of firm. Our results therefore suggest only a few firms within business related services used the COVID-induced shift to WFH for upskilling of their workforce in 2020. Most other firms only increased their demand for manual tasks instead.

While we view our findings credible and intuitive in light of the unique features of COVID-19, they also reveal important differences to other countries. Notably, Blanas & Oikonomou (2022) document an increase in demand for occupations intensive in analytic and interactive tasks in the US. Yet, unlike our study, Blanas & Oikonomou (2022) study state-level variation, possibly masking underlying heterogeneity at the local labor market level. Moreover, we believe part of this discrepancy can be attributed to different strategies in combatting

COVID-19. While the US experienced high unemployment rates of up to 15%, Germany relied heavily on a short-time work model ("Kurzarbeit"), thereby capping unemployment at 6%. Based on the experiences from the Great Recession 2007/08, upskilling (in form of higher analytic and interactive task requirements) is more likely to occur in regions with high unemployment rates (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019). Therefore, COVID-induced regional variation in unemployment was *less* pronounced in Germany, suppressing upskilling incentives. This example demonstrates the importance of cross-country comparisons to understand how different policies affected the recovery process of the pandemic —an important avenue for future research.

More broadly, we suspect the pandemic-induced changes in task demand have the potential to reinforce structural differences. Prior to the pandemic, differences in task demand across local labor markets decreased, suggesting a convergence in task demand. Since the pandemic, however, differences in task demand across local labor markets have increased, indicating a divergence. Unfortunately, we are not able to explain yet whether this divergence is lasting. While we do provide various important descriptive statistics until December 2021, our current identification strategy only allows us to draw causal inference until December 2020.

In an upcoming revise of our draft we will thus extend the scope of our paper as follows. First, we will be implementing a continuous treatment strategy (rather than binary treatment) to exploit differential treatment intensities across LLMs more explicitly. As COVID-19 affected all regions and firms, there are no natural control groups. To provide remedy, we are employing propensity score weighting methods to control for systematic differences pre-pandemic, similar to Ben Yahmed, Berlingieri & Brüll (2022). This method has the key advantage that it allows us to supplement our current (short-term) causal analysis with medium-term causal effects. For this task, we will take advantage of continuous data updates, allowing us to study changes in task demand until December 2022. Second, we seek to improve our understanding about the underlying mechanisms of our results. To this end, we will be incorporating more data on (i) ICT-related technology differences across firms and local labor markets and (ii) institutional differences via differential propensities to employ short-time work.

References

Adams-Prassl, A., Boneva, T., Golin, M. & Rauh, C. (2020), 'Inequality in the impact of the coronavirus shock: Evidence from real time surveys', *Journal of public economics* **189**, 104245. 1

Adams-Prassl, A., Boneva, T., Golin, M. & Rauh, C. (2022), 'Work that can be done from home: evidence on variation within and across occupations and industries', *Labour Economics* **74**, 102083. 1, 3, 14

Alipour, J.-V., Falck, O. & Schüller, S. (2020), 'Germany's capacities to work from home', *IZA DP No. 13152* . 1, 3, 14

Alipour, J.-V., Langer, C. & O'Kane, L. (2021), 'Is working from home here to stay? a look at 35 million job ads', (22), 44–46. 1, 3

Antonczyk, D., Fitzenberger, D. & Leuschner, U. (2009), 'Can a task-based approach explain the recent changes in the german wage structure?', *Jahrbücher für Nationalökonomie und Statistik* (229), 214–238. 8

Arthur, R. (2021), 'Studying the uk job market during the covid-19 crisis with online job ads', *PloS one* **16**(5), e0251431. 2, 11, 17, 42

Atalay, E., Phongthiengtham, P., Sotelo, S. & Tannenbaum, D. (2018), 'New technologies and the labor market', *Journal of Monetary Economics* **97**, 48–67. 1, 3

Autor, D. H., Dorn, D. & Hanson, G. H. (2015), 'Untangling trade and technology: Evidence from local labour markets', *The Economic Journal* **125**(584), 621–646. 1

Autor, D. H. & Handel, M. J. (2013), 'Putting tasks to the test: Human capital, job tasks, and wages', *Journal of Labor Economics* **31**(S1), S59–S96. 3

Autor, D. H., Levy, F. & Murnane, R. J. (2003), 'The skill content of recent technological change: An empirical exploration', *The Quarterly Journal of Economics* **118**(4), 1279–1333. 7

Bachmann, R., Cim, M. & Green, C. (2019), 'Long–run patterns of labour market polarization: Evidence from german micro data', *British Journal of Industrial Relations* **57**(2), 350–376. 1, 3

Bachmann, R., Demir, G. & Frings, H. (2021), 'Labor market polarization, job tasks and monopsony power', *Journal of Human Resources* pp. 0219–10011R1. 3, 8

Bachmann, R., Demir, G., Green, C. & Uhlrendorff, A. (2022), 'The role of within-occupation task changes in wage development', *Ruhr Economic Papers No. 975* . 3

Bamieh, O. & Ziegler, L. (2020), 'How does the covid-19 crisis affect labor demand? an analysis using job board data from austria', *IZA DP No. 13801* . 1, 3, 42

Barrero, J. M., Bloom, N. & Davis, S. (2020), 'Covid-19 is also a reallocation shock'. 1, 3

Barrero, J. M., Bloom, N. & Davis, S. (2021), 'Why working from home will stick'. 1, 3

Ben Yahmed, S., Berlingieri, F. & Brüll, E. (2022), 'Adjustments of local labour markets to the covid-19 crisis: the role of digitalisation and working-from-home', *ZEW DISCUSSION PAPER No. 22-031* . 1, 3, 15, 16, 18

Blanas, S. & Oikonomou, R. (2022), 'Civid-induced economic uncertainty, tasks, and occupational demand', *SSRN Journal* . 3, 4, 6, 11, 17

Blau, F. D., Koebe, J. & Meyerhofer, P. A. (2021), 'Who are the essential and frontline workers?', *Business economics (Cleveland, Ohio)* **56**(3), 168–178. 2, 11, 17

Blien, U., Dauth, W. & Roth, D. H. (2021), 'Occupational routine intensity and the costs of job loss: evidence from mass layoffs', *Labour Economics* **68**, 101953. 1

Campello, M., Kankanhalli, G. & Muthukrishnan, P. (2020), 'Corporate hiring under covid-19: Labor market concentration, downskilling, and income inequality'. 1, 3

Carvalho, B. P., Peralta, S. & Pereira Dos Santos, J. (2021), 'Regional and sectorial impacts of the covid-19 crisis: Evidence from electronic payments', *Journal of regional science* . 3, 8, 9

Cassidy, H. (2017), 'Task variation within occupations', *Industrial Relations: A Journal of Economy and Society* **56**(3), 393–410. 3

Cortes, G. M., Jaimovich, N. & Siu, H. (2021), 'The growing importance of social tasks in high-paying occupations: Implications for sorting'. 3

Cortes, G. M., Lerche, A., Schoenberg, U. & Tschopp, J. (2021), 'Technological change, firm heterogeneity and wage inequality'. 3

Dauth, W., Findeisen, S., Suedekum, J. & Woessner, N. (2021), 'The adjustment of labor markets to robots', *Journal of the European Economic Association* **19**(6), 3104–3153. 6

de Fraja, G., Matheson, J. & Rockey, J. (2021), 'Zoomshock : the geography and local labour market consequences of working from home', *Covid Economics* **64**(64), 1–41. 3

Deming, D. & Kahn, L. B. (2018), 'Skill requirements across firms and labor markets: Evidence from job postings for professionals', *Journal of Labor Economics* **36**(S1), S337–S369. 4

Dengler, K., Matthes, B. & Paulus, W. (2014), 'Occupational tasks in the german labour market: An alternative measurement on the basis of an expert database', *FDZ Methodenreport 201412* . 7

Dingel, J. I. & Neiman, B. (2020), 'How many jobs can be done at home?', *Journal of public economics* **189**, 104235. 1, 3, 14

Dustmann, C., Lindner, A., Schönberg, U., Umkehrer, M. & vom Berge, P. (2021), 'Reallocation effects of the minimum wage', *The Quarterly Journal of Economics* **137**(1), 267–328. 6

Forsythe, E. (2022), 'Why don't firms hire young workers during recessions?', *The Economic Journal* **132**(645), 1765–1789. 4

Forsythe, E., Kahn, L. B., Lange, F. & Wiczer, D. (2020*a*), 'Labor demand in the time of covid-19: Evidence from vacancy postings and ui claims', *Journal of public economics* **189**, 104238. 1, 3, 42

Forsythe, E., Kahn, L., Lange, F. & Wiczer, D. (2020*b*), 'Searching, recalls, and tightness: An interim report on the covid labor market'. 1, 3

Gathmann, C. & Schönberg, U. (2010), 'How general is human capital? a task–based approach', *Journal of Labor Economics* **28**(1), 1–49. 3, 7

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–574. 6

Goecke, H. & Rusche, C. (2022), 'Coronapandemie: Auswirkungen auf den einzelhandel in den innenstädten', *Wirtschaftsdienst* **102**(4), 298–302. 11

Gürtzgen, N., Lochner, B., Pohlan, L. & van den Berg, G. J. (2021), 'Does online search improve the match quality of new hires?', *Labour Economics* **70**, 101981. 41

Hamann, S., Niebuhr, A., Roth, D. & Sieglen, G. (2022), 'How does the covid-19 pandemic affect regional labour markets and why do large cities suffer most?', **mimeo**. 12

Hensvik, L., Le Barbanchon, T. & Rathelot, R. (2021), 'Job search during the covid-19 crisis', *Journal of public economics* **194**, 104349. 1, 3, 42

Hershbein, B. & Kahn, L. B. (2018), 'Do recessions accelerate routine-biased technological change? evidence from vacancy postings', *American Economic Review* **108**(7), 1737–1772. 4, 12, 16, 18, 41, 42

Hirsch, B., Jahn, E. J., Manning, A. & Oberfichtner, M. (2022), 'The urban wage premium in imperfect labor markets', *Journal of Human Resources* **57**(S), S111–S136. 6

Hoseini, M. & Valizadeh, A. (2021), 'The effect of covid-19 lockdown and the subsequent reopening on consumption in iran', *Review of economics of the household* **19**(2), 373–397. 3, 8

Kosfeld, R. & Werner, A. (2012), 'Deutsche arbeitsmarktregionen – neuabgrenzung nach den kreisgebietsreformen 2007–2011', *Spatial Research and Planning* **70**(1). 6

Marinescu, I. E., Skandalis, D. & Zhao, D. (2020), 'Job search, job posting and unemployment insurance during the covid-19 crisis', *SSRN Electronic Journal* . 3

Modestino, A. S., Shoag, D. & Ballance, J. (2019), 'Upskilling: Do employers demand greater skill when workers are plentiful?', *The Review of Economics and Statistics* **4**, 1–46. 4, 12, 16, 18, 42

Regionalstatistik (2022/10/25), 'Regionaldatenbank deutschland'.
URL: *https://www.regionalstatistik.de/genesis/online/* 6

Rengers, M. (2018), 'Internetgestützte erfassung offener stellen: Machbarkeitsstudie im rahmen eines essnet-projekts zu big data', *WISTA - Wirtschaft und Statistik* **5**. 41

Rohrbach-Schmidt, D. (2019), 'Putting tasks to the test: The case of germany', *Social Inclusion* **7**(3), 122–135. 3

Rohrbach-Schmidt, D. & Tiemann, M. (2013), 'Changes in workplace tasks in Germany—

evaluating skill and task measures', *Journal for Labour Market Research* **46**(3), 215–237. 7

Shen, K. & Taska, B. (2020), 'Measuring the impacts of covid-19 on job postings in australia using a reweighting-estimation-transformation approach', *Australian Journal of Labour Economics* **23**(2), 153–171. 3, 42

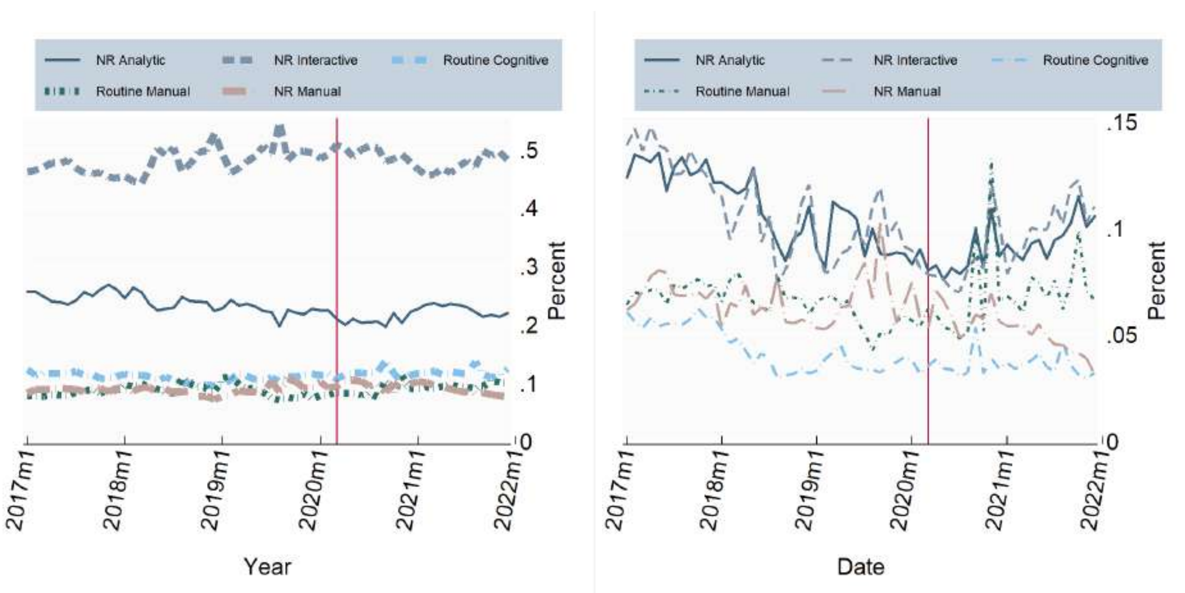
Spitz-Oener, A. (2006), 'Technical change, job tasks, and rising educational demands: Looking outside the wage structure', *Journal of Labor Economics* **24**(2), 235–270. 3, 7

Stops, M., Bächmann, A.-C., Glassner, R., Janser, M., Matthes, B., Metzger, L.-J. & Müller, Christoph, Seitz, Joachim (2021), 'Extracting skill requirements from job ads - the machbarkeitsstudie kompetenz-kompass', *IAB-Forschungsbericht No. 7/2021* . 5

Storm, E. (2022*a*), 'On the measurement of tasks: Does expert data get it right?', *Ruhr Economic Papers, No. 948* . 3, 7, 8

Storm, E. (2022*b*), 'Task specialization and the native–foreign wage gap', *LABOUR* **36**(2), 167–195. 3, 7, 8

A Descriptive Figures



(a) Task Demand (Aggregate) (b) Task Demand (90-10 Differential)

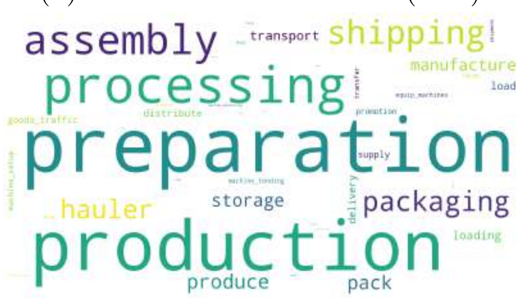
NOTE. —Panel (a) displays aggregate task demand across 141 German local labor markets (LLM). Panel (b) displays the average 90-10 differential of task demand across all LLMs for each date, capturing between-LLM differences.

Figure 1: Demand for Tasks in Germany, 2017-01 - 2021-12

(a) Non-Routine Analytic (NRA)

(b) Non-Routine Interactive (NRI)

(c) Routine Cognitive (RC)

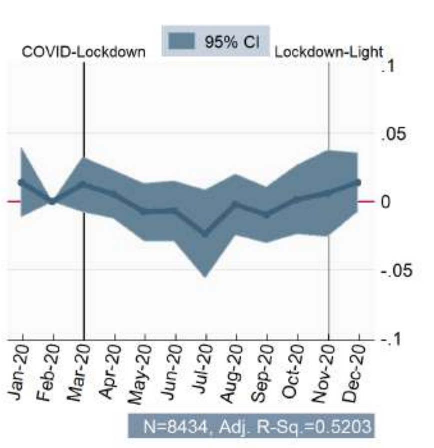


(d) Routine Manual (RM)

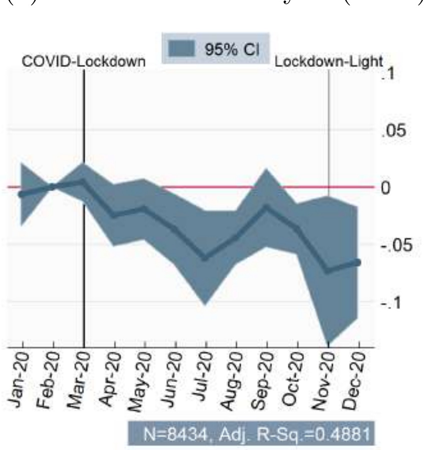
(e) Non-Routine Manual (NRM)

Figure 2: Word clouds of job activities within task groups

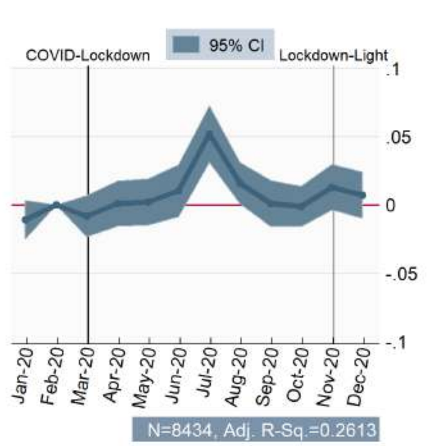
B Results: Local Labor Market level



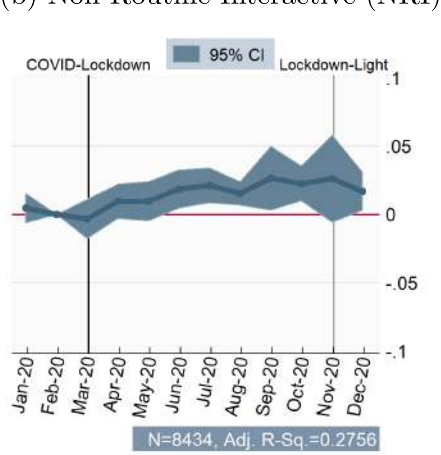
(a) Non-Routine Analytic (NRA)



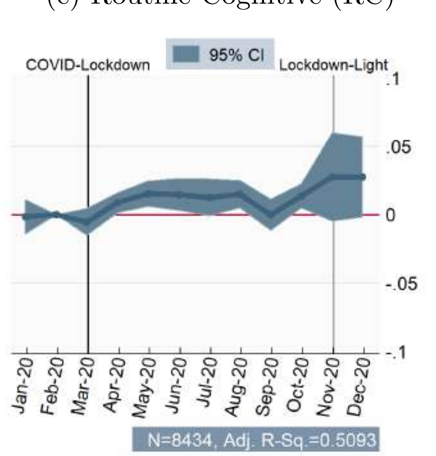
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



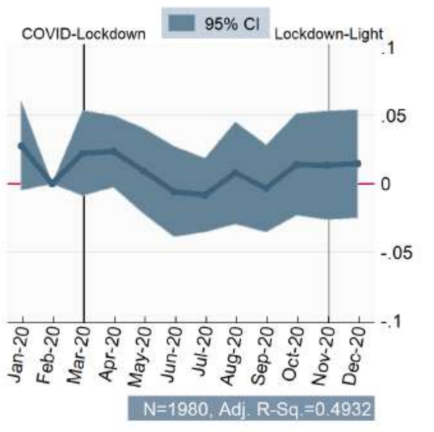
(d) Routine Manual (RM)



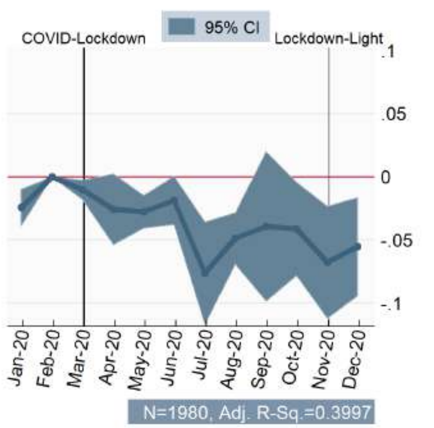
(e) Non-Routine Manual (NRM)

NOTE. —Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

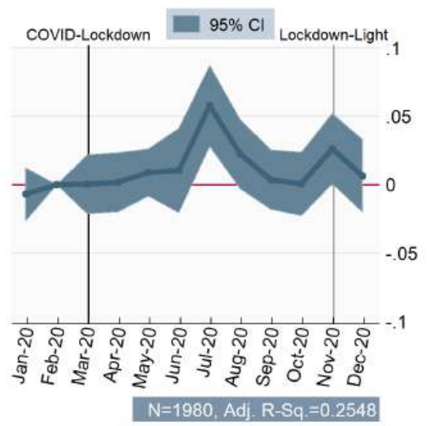
Figure 3: DiD Estimates, 2020-01 - 2020-12



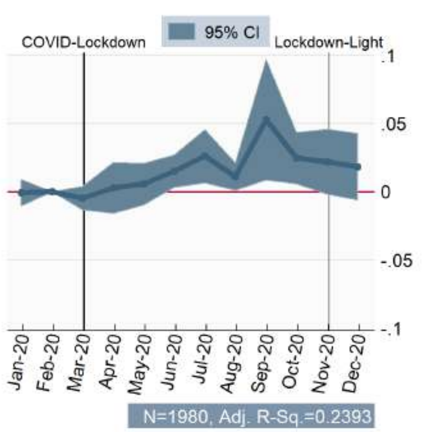
(a) NR Analytic (Agglomeration)



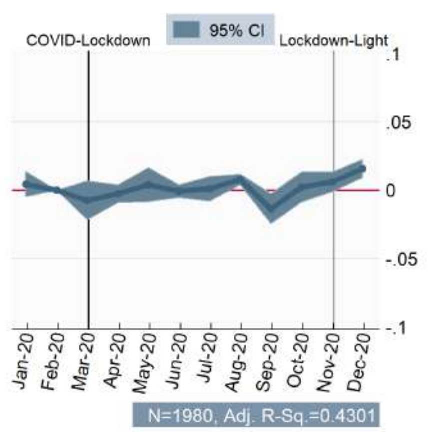
(b) NR Interactive (Agglomeration)



(c) Routine Cognitive (Agglomeration)



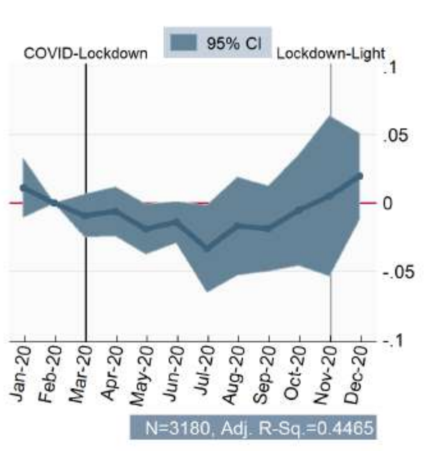
(d) Routine Manual (Agglomeration)



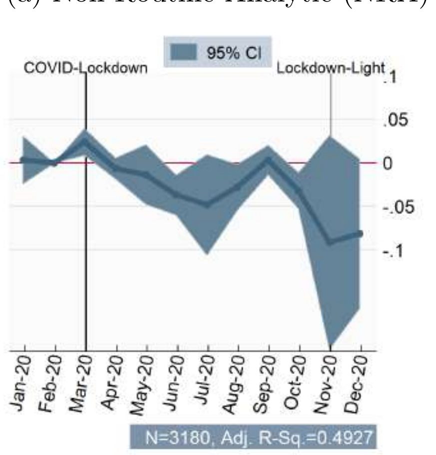
(e) N Manual (Agglomeration)

NOTE. —Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

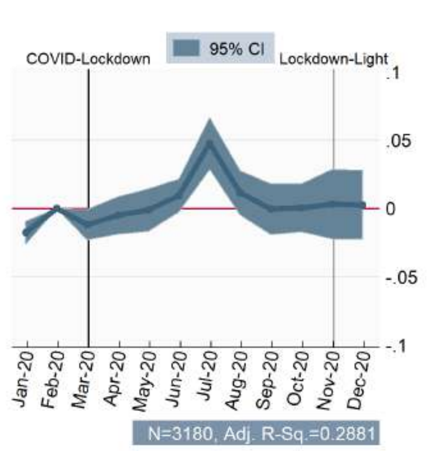
Figure 4: DiD Estimates, 2020-01 - 2020-12, Agglomeration areas



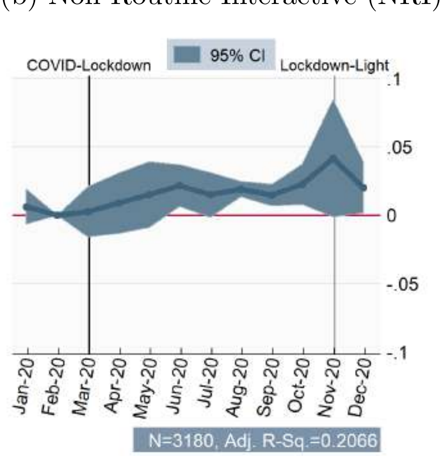
(a) Non-Routine Analytic (NRA)



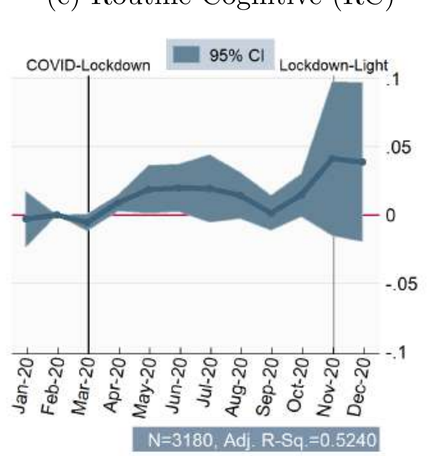
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



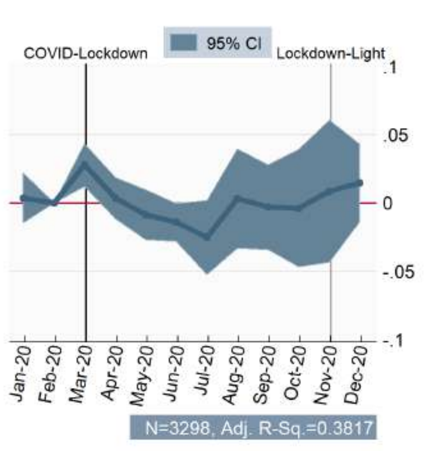
(d) Routine Manual (RM)



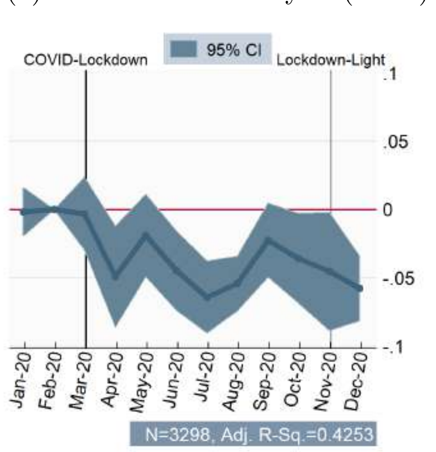
(e) Non-Routine Manual (NRM)

NOTE. —Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

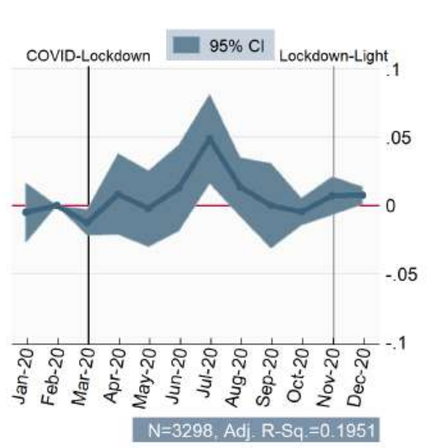
Figure 5: DiD Estimates, 2020-01 - 2020-12, Urbanized areas



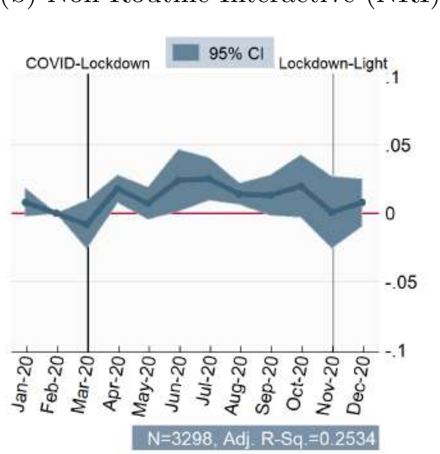
(a) Non-Routine Analytic (NRA)



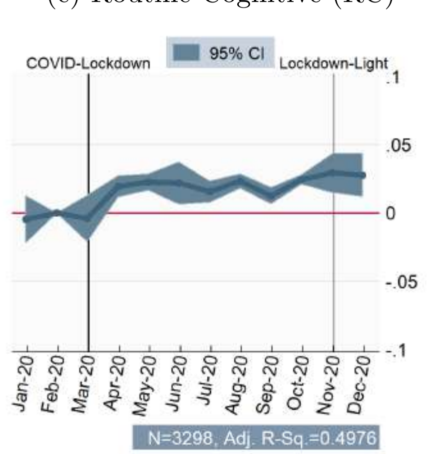
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



(d) Routine Manual (RM)

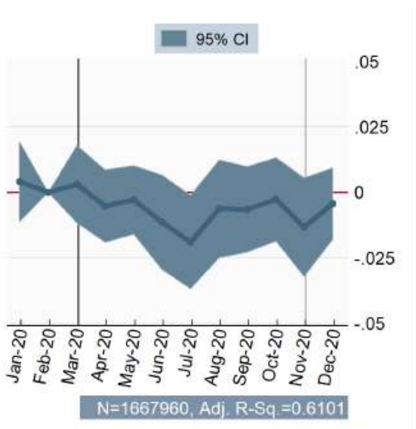


(e) Non-Routine Manual (NRM)

NOTE. —Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

Figure 6: DiD Estimates, 2020-01 - 2020-12, Rural areas

C Results: Firm-level



(a) Non-Routine Analytic (NRA)

(b) Non-Routine Interactive (NRI)

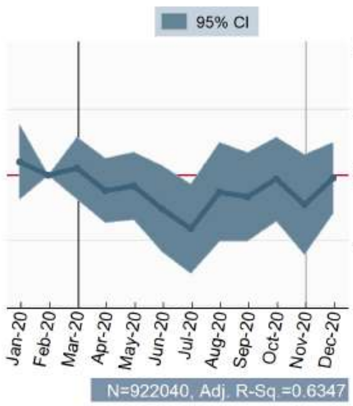
(c) Routine Cognitive (RC)

(d) Routine Manual (RM)

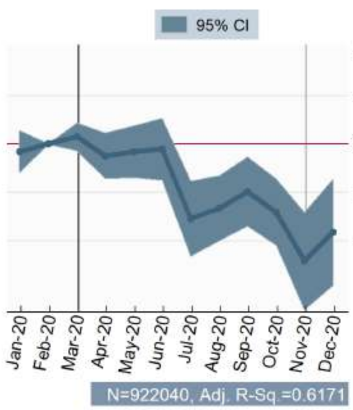
(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

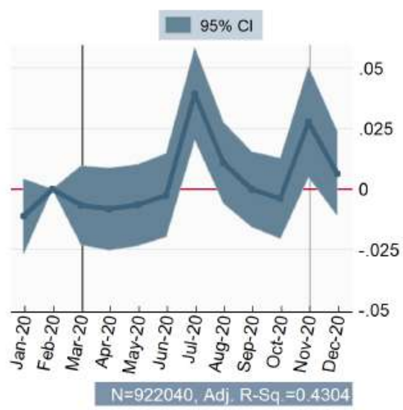
Figure 7: Difference-in-Difference Estimates (Baseline): Firm-level, 2020-01 - 2020-12



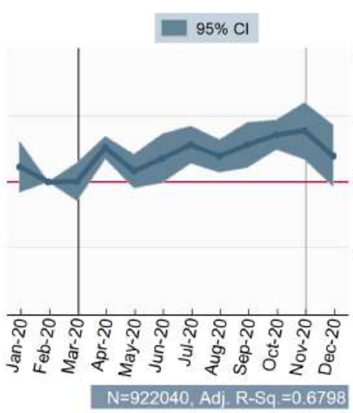
(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

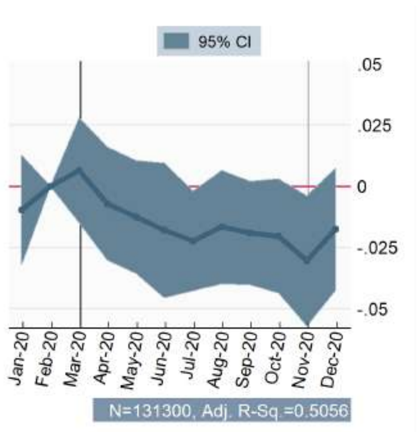


(d) Routine Manual (RM)

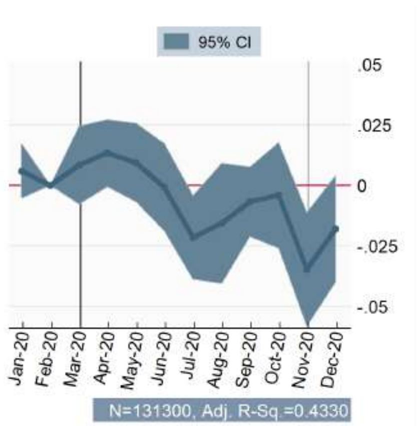
(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 8: Difference-in-Difference Estimates: Firm-level (*Large Firms*), 2020-01 - 2020-12



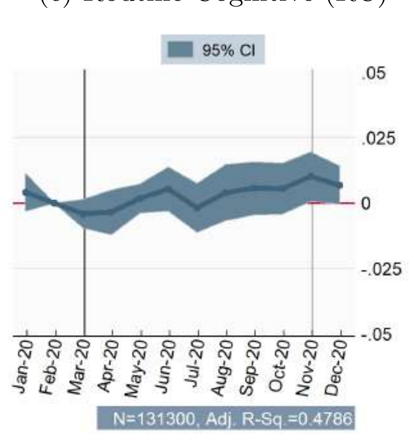
(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)

(c) Routine Cognitive (RC)

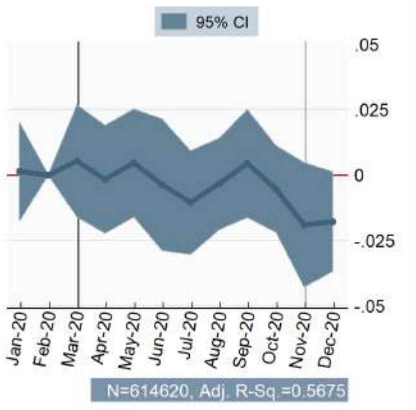
(d) Routine Manual (RM)



(e) Non-Routine Manual (NRM)

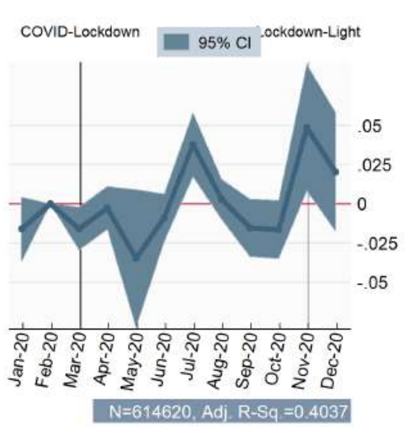
NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 9: Difference-in-Difference Estimates: Firm-level (*Mid-sized Firms*), 2020-01 - 2020-12



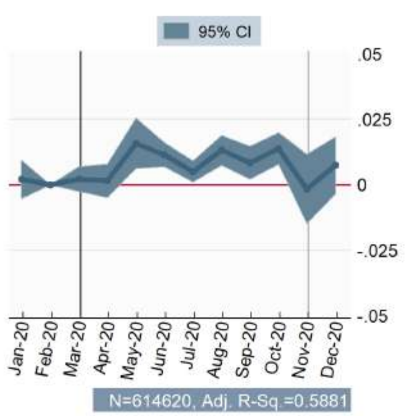
(a) Non-Routine Analytic (NRA)

(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

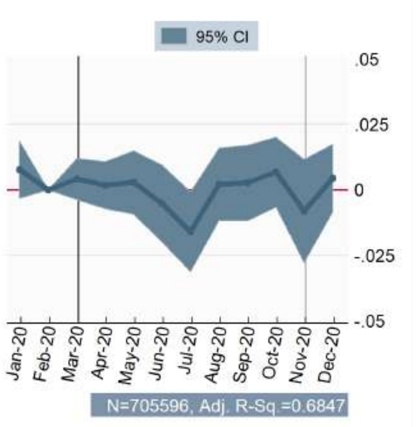
(d) Routine Manual (RM)



(e) Non-Routine Manual (NRM)

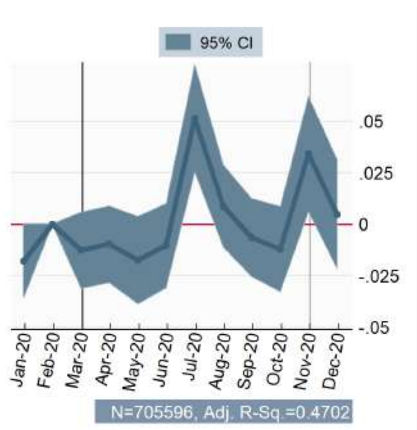
NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 10: Difference-in-Difference Estimates: Firm-level (*Small Firms*), 2020-01 - 2020-12

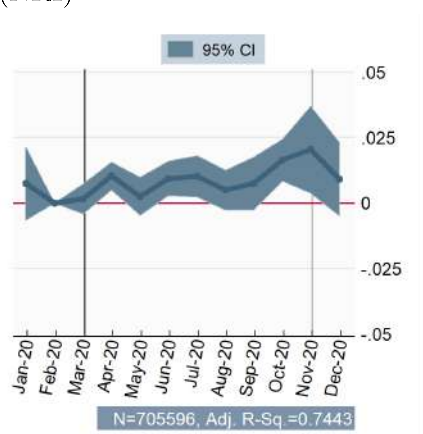


(a) Non-Routine Analytic (NRA)

(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



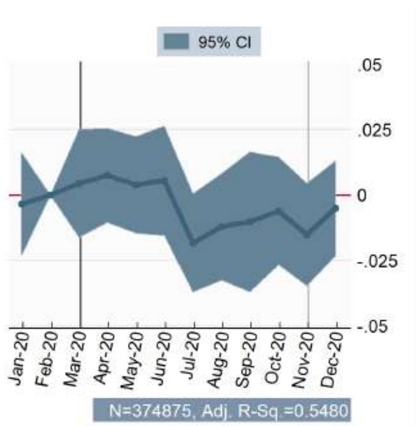
(d) Routine Manual (RM)

(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 11: Difference-in-Difference Estimates: Firm-level (*business related services*), 2020-01 - 2020-12

(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)

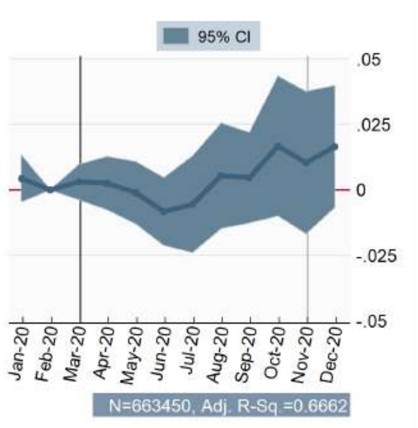
(c) Routine Cognitive (RC)

(d) Routine Manual (RM)

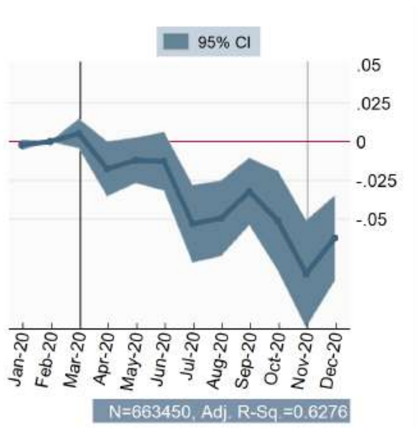
(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

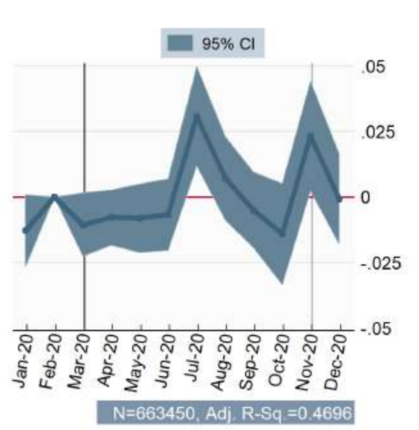
Figure 12: Difference-in-Difference Estimates: Firm-level (*personal services*), 2020-01 - 2020-12



(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)



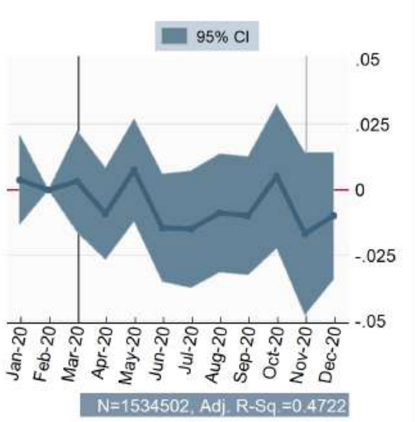
(c) Routine Cognitive (RC)

(d) Routine Manual (RM)

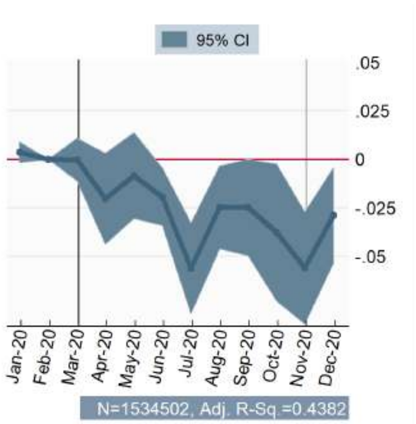
(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

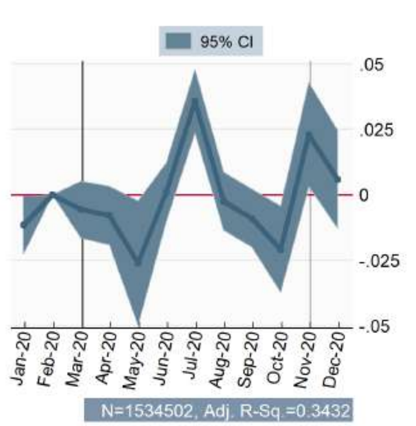
Figure 13: Difference-in-Difference Estimates: Firm-level (*Firms offering WFH pre-COVID*), 2020-01 - 2020-12



(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

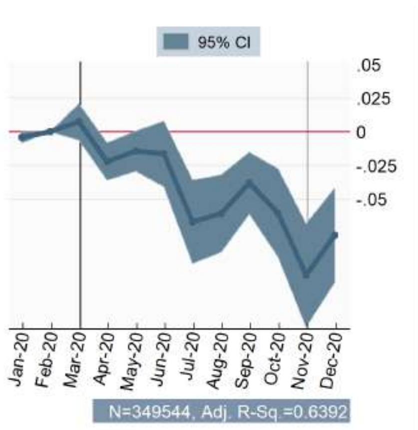
(d) Routine Manual (RM)

(e) Non-Routine Manual (NRM)

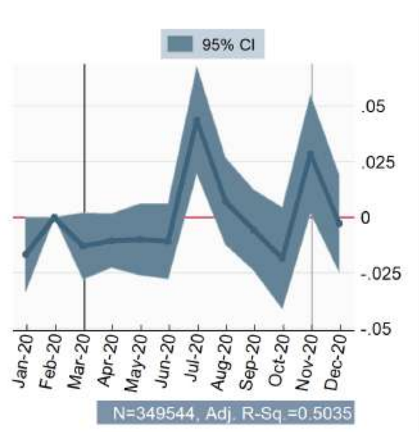
NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 14: Difference-in-Difference Estimates: Firm-level (*Firms* ***not*** *offering WFH pre-COVID*), 2020-01 - 2020-12

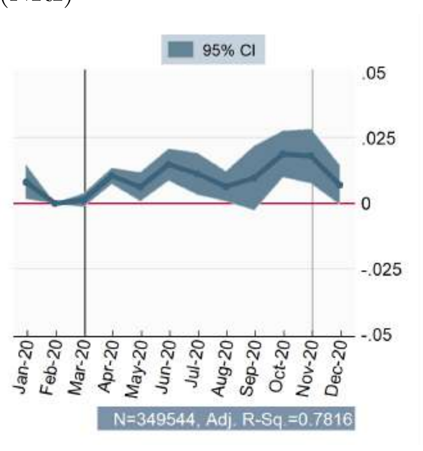
(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

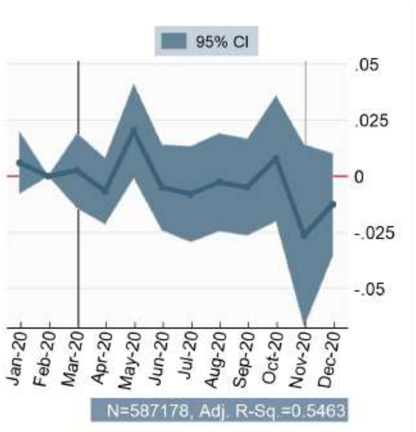


(d) Routine Manual (RM)

(e) Non-Routine Manual (NRM)

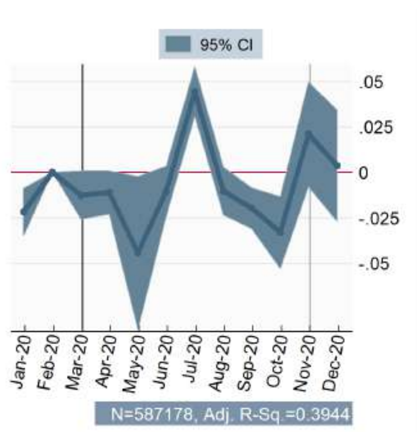
NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 15: Difference-in-Difference Estimates: Firm-level (*business related services, WFH Firms*), 2020-01 - 2020-12

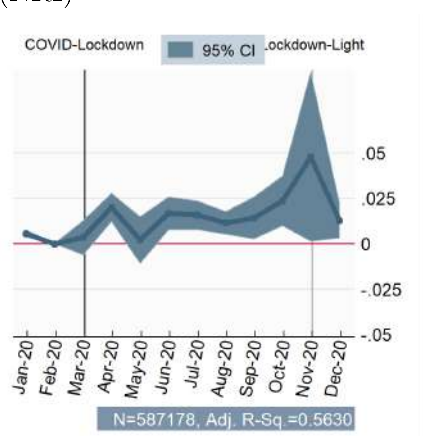


(a) Non-Routine Analytic (NRA)

(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

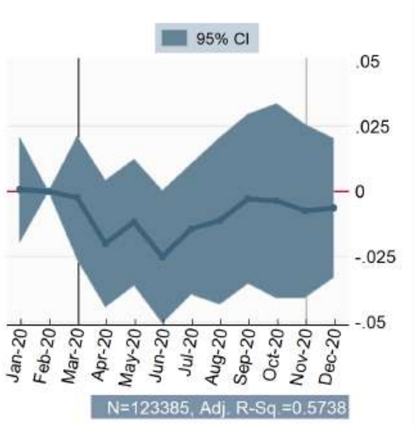


(d) Routine Manual (RM)

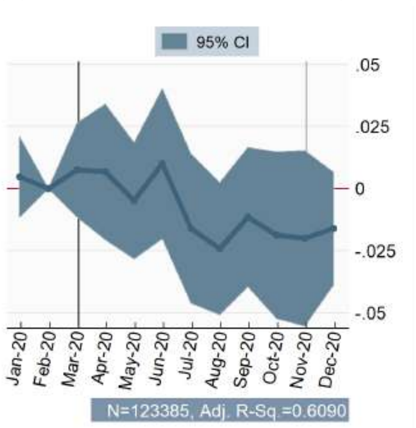
(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

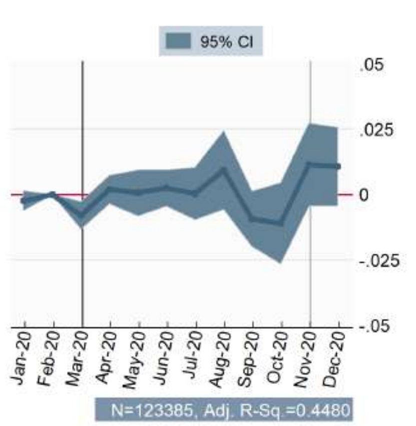
Figure 16: Difference-in-Difference Estimates: Firm-level (*business related services, non-WFH Firms*), 2020-01 - 2020-12



(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)

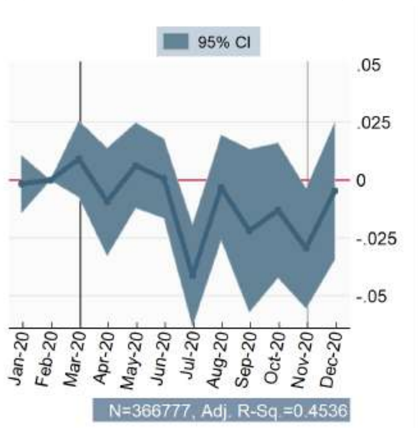
(d) Routine Manual (RM)

(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

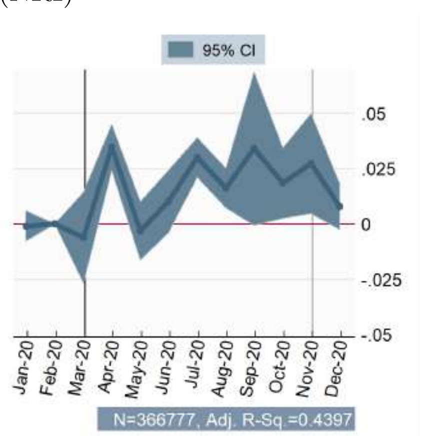
Figure 17: Difference-in-Difference Estimates: Firm-level (*personal services, WFH Firms*), 2020-01 - 2020-12

(a) Non-Routine Analytic (NRA)



(b) Non-Routine Interactive (NRI)

(c) Routine Cognitive (RC)



(d) Routine Manual (RM)

(e) Non-Routine Manual (NRM)

NOTE. —Point estimates are displayed with a 95% Confidence Interval.

Figure 18: Difference-in-Difference Estimates: Firm-level (*business related services, non-WFH Firms*), 2020-01 - 2020-12

D Appendix

D.1 External Validity of OJV Data

Figure 19a shows the number of OJV over time by source platforms. Overall, we see an increasing trend of the number of postings over time. In principle, this pattern can be explained by two factors. First, an increasing trend over time, i.e., firms may use their websites and job boards more often to post jobs online. Second, methodological changes, e.g., our private partner updates its scraping method and thus adds more sources. Rising levels of digitalization and the growing popularity of online job search by job seekers likely contribute to the increasing trend in OJV. We further find evidence that methodological changes matter as well since the composition of source platforms has changed over time. While (fee paying) job boards represented about 50% of all postings in 2017, their share increased to 70% by the end of 2021. This increase has come primarily at the expense of headhunters whose share decreased from 17% to less than 2% during the same time. These compositional changes demonstrate the need to validate the representativeness of OJV data.

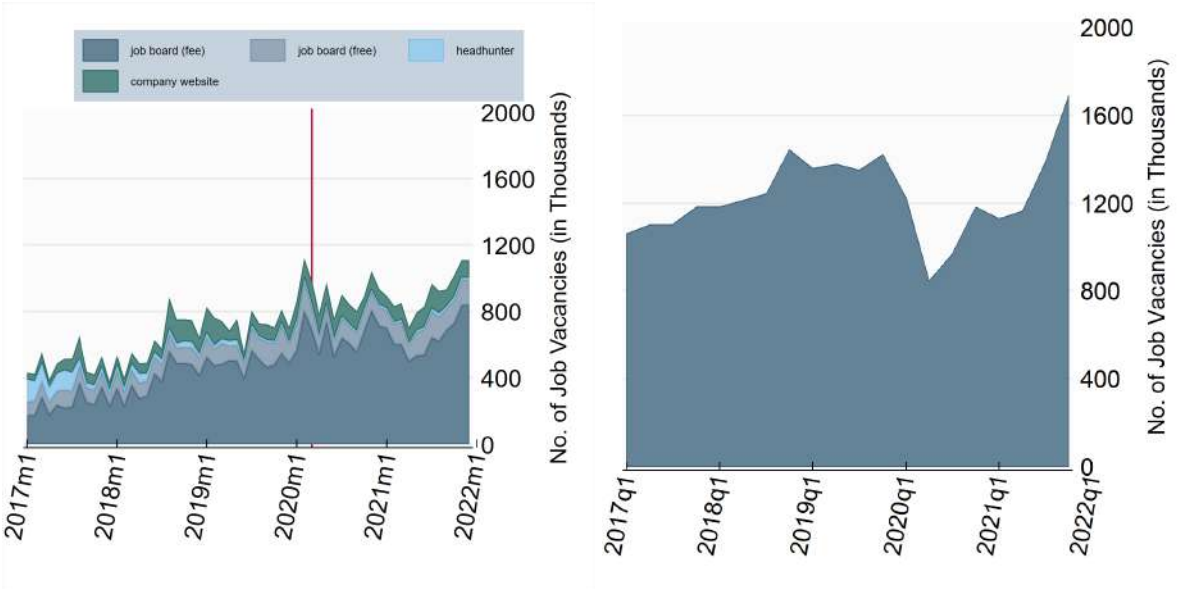
[Figure 19 here]

We follow common practice in the literature by comparing our OJV data with representative information on vacancies from official sources (Hershbein & Kahn 2018, Rengers 2018). Hershbein & Kahn (2018) compare characteristics of the job postings from Lightcast (formerly Burning Glass Technologies) with the Bureau of Labor Statistics' Job Openings and Labor Market Turnover (JOLTS) survey and other data sources for the US at the aggregate level and by industries. Likewise, Rengers (2018) makes similar comparisons for Germany with data from the Federal Employment Agency (BA) and the IAB Job Vacancy Survey. Especially relevant for our purposes, the IAB Job Vacancy Survey is a representative survey and measures the aggregate labor demand and the recruiting behavior of firms in Germany since 1989, making it a well-suited survey for the analysis of recruitment processes (Gürtzgen, Lochner, Pohlan & van den Berg 2021). Below, we address these concerns by first studying aggregate trends and subsequently breaking down our OJV data by industries.

First, Figure 19 compares the (aggregate) evolution of vacancies taken from the IAB Job Vacancy Survey from 2017Q1 - 2021Q4 (2021 values are estimates) with our OJV data. Note that the IAB data reflects stock information, while our data is a measure for inflows of job postings. Despite these methodological differenes, the two graphs display similar trends. Both display a steady increase in postings from 2017 until early 2020 with a sharp decrease at the onset of the pandemic in March 2020. While the stock of vacancies decreased by 40% between 2019Q4 and 2020Q2 based on the IAB Vacancy Panel, the inflows of vacancies in our OJV data decreased by 30% from December 2019 until June 2020. Both time series display a sharp subsequent rebound, leading to a catch-up to pre-COVID vacancy levels by the end of 2020. Moreover, the magnitude of the drop and rebound in job vacancies during the pandemic is consistent with previous findings in the literature from comparable countries, such as Australia (Shen & Taska 2020), Austria (Bamieh & Ziegler 2020), Sweden (Hensvik, Le Barbanchon & Rathelot 2021), the UK (Arthur 2021), and the US (Forsythe, Kahn, Lange & Wiczer 2020*a*). Hence, both, the cyclicality of job postings and the magnitude in collapse and recovery of postings, lend credence to the validity of our data.

[Figure 20 here]

Second, we divide our vacancies into six broad industries for ease of exposition: (i) manufacturing, (ii) retail & hospitality, (iii) information & communication, (iv) professional services, (v) personal services, and (vi) other industries. Figure 20 summarizes this industrial breakdown and provides three key takeaways. First, all industries are covered and well-represented in our data. Second, service industries, comprising professional and personal services, are the most important industry groups. On average, these broad industries comprise around half of all vacancies. Third, the industry composition in our data has become more balanced over time. While the share of services decreased from 60% to 45% from 2017 until 2021, manufacturing and retail & hospitality have experienced rising coverage (in each industry from 15% to 20%). We interpret these developments favorably as the descriptive statistics support the quality of our data and its broadly representative nature. Part of this takeaway is attributed to the fact that our data begins in 2017. Internet access and especially online job search have already been common at this point, a distinguishing feature from the earliest possible OJV data in the US in the mid 2000s, a time during which online job posting was concentrated among professionals (Hershbein & Kahn 2018, Modestino, Shoag & Ballance 2019).



(a) OJV data, by source (Inflow) (b) IAB Vacancy Panel (Stock)

NOTE. —Panel 19a displays the number of online job vacancies that are posted each month in our data, i.e., monthly inflows, broken down by the type of source platform. Panel 19b displays the stock of vacancies firms report to the IAB for each quarter. The values for 2021Q1 onward are estimates as final numbers are not available yet.

Figure 19: Number of Online Job Vacancies over Time, 2017-01 - 2021-12

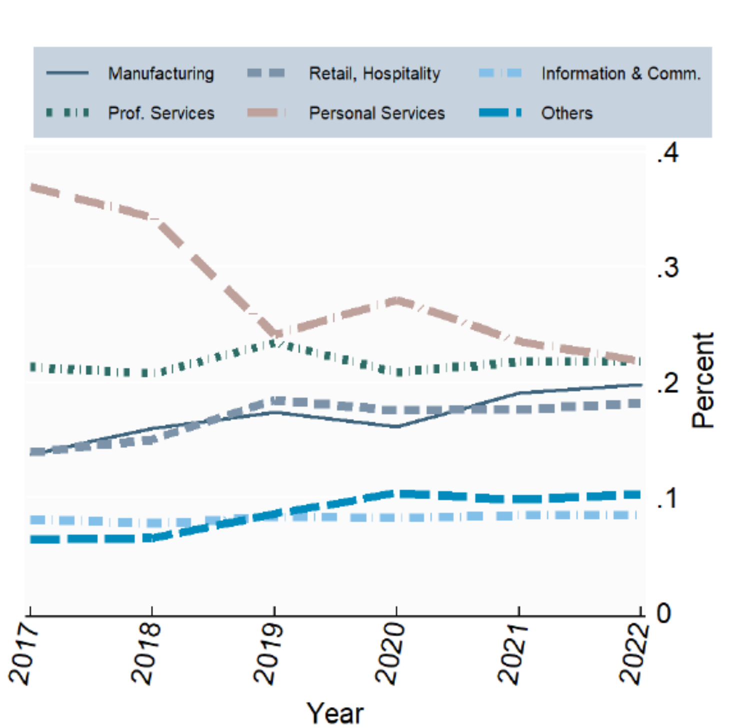
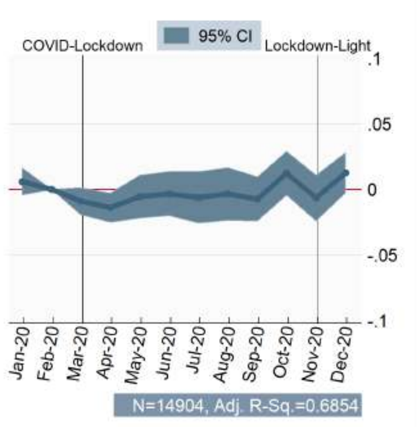
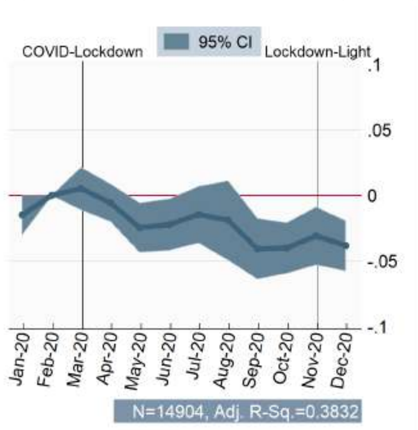


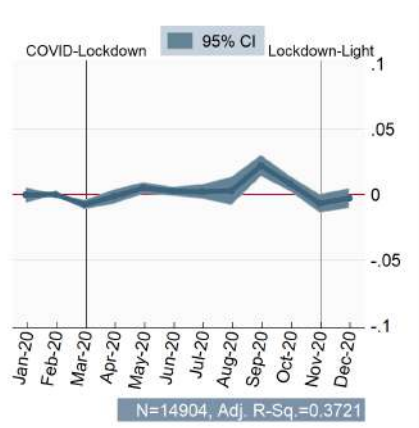
Figure 20: Industry Composition of Online Job Vacancies year, 2017 - 2021



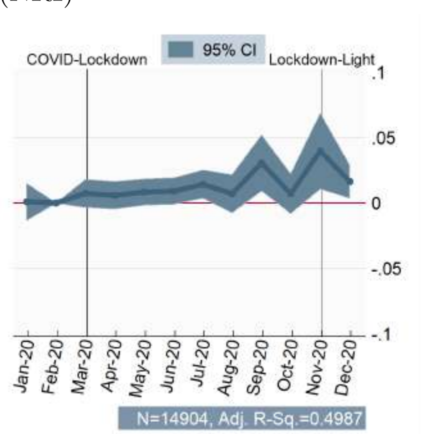
(a) Non-Routine Analytic (NRA)



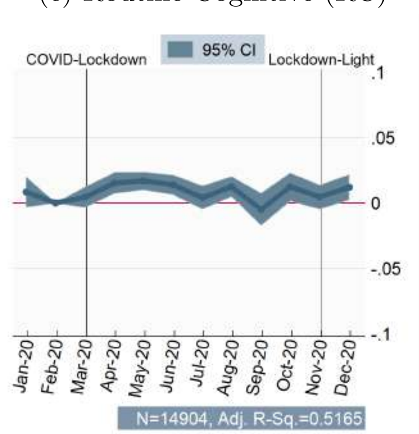
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



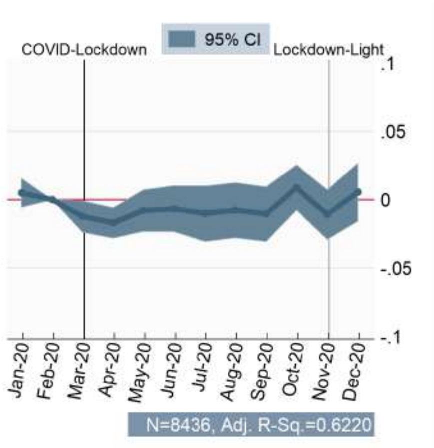
(d) Routine Manual (RM)



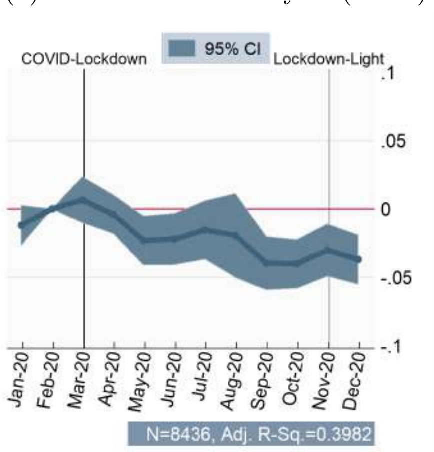
(e) Non-Routine Manual (NRM)

NOTE. —These robustness tests have a more narrow definition of LLMs, comprising 258 rather than 141 LLMs. Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

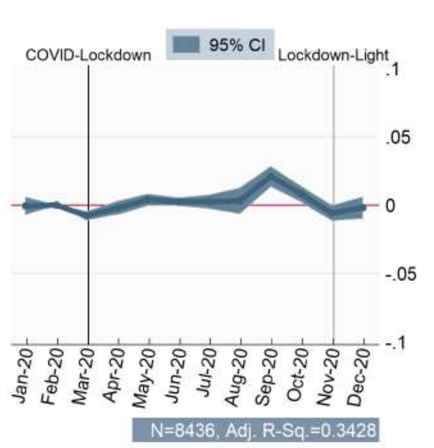
Figure 21: DiD Estimates: LLM-level (*Alternative LLM Definition*), 2020-01 - 2020-12



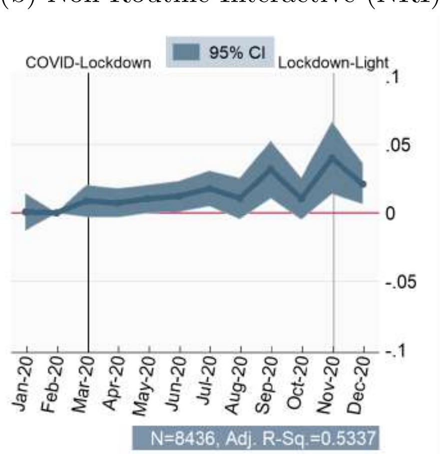
(a) Non-Routine Analytic (NRA)



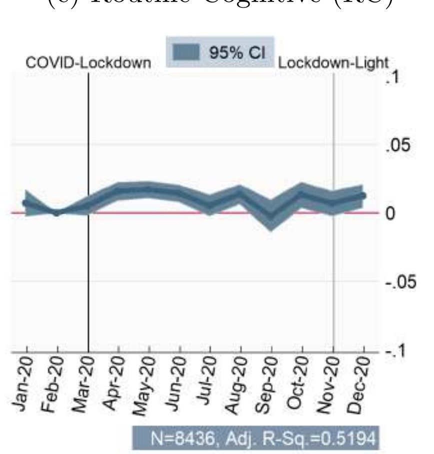
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



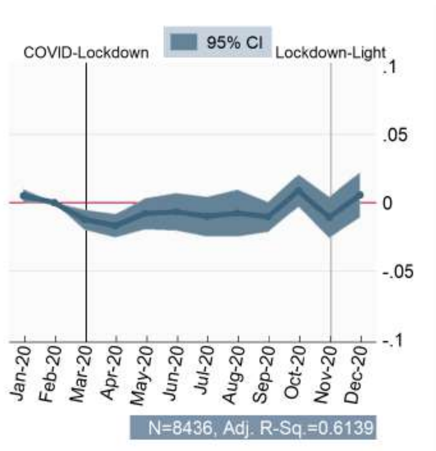
(d) Routine Manual (RM)



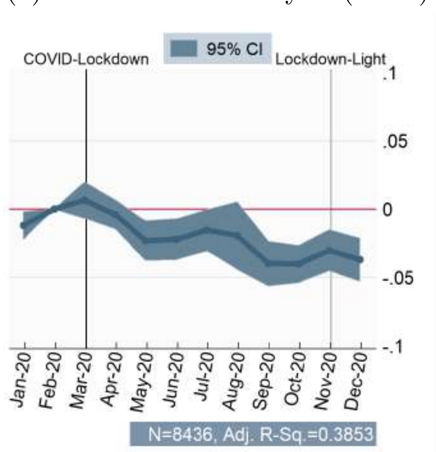
(e) Non-Routine Manual (NRM)

NOTE. —These robustness tests have month × LLM FE rather than month and LLM. Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

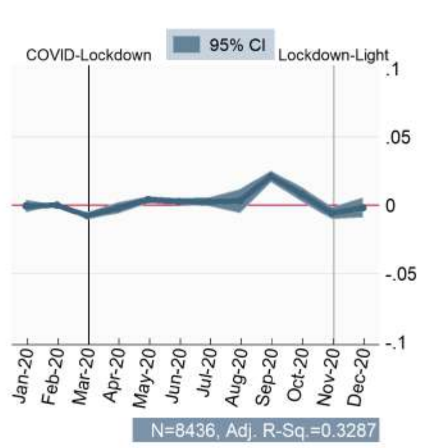
Figure 22: DiD Estimates: LLM-level (*Changing FE*), 2020-01 - 2020-12



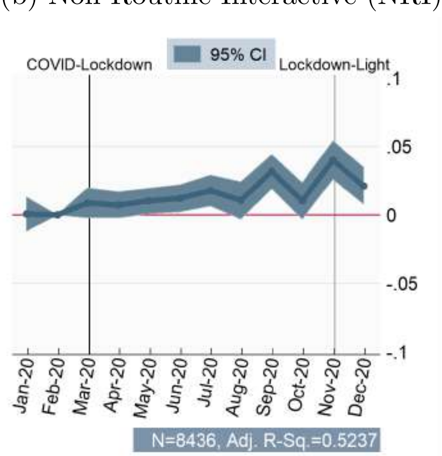
(a) Non-Routine Analytic (NRA)



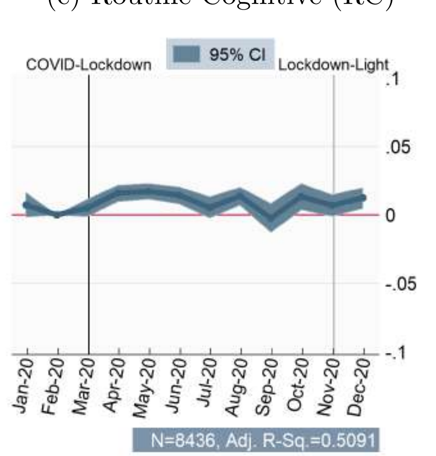
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



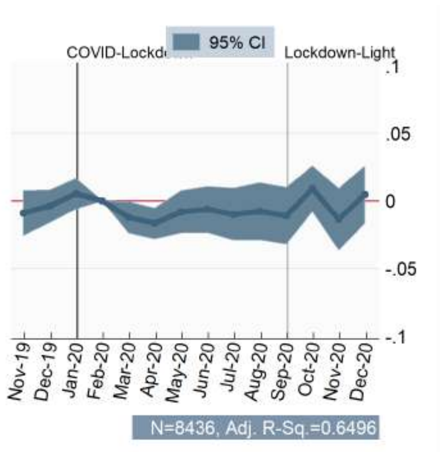
(d) Routine Manual (RM)



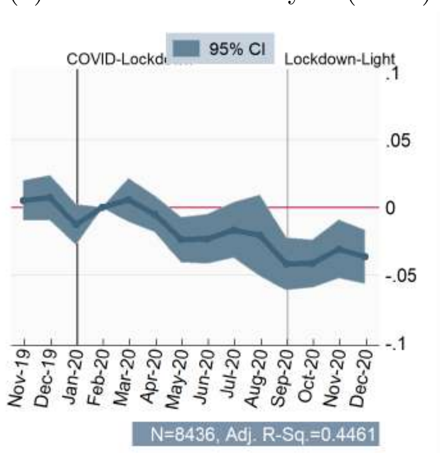
(e) Non-Routine Manual (NRM)

NOTE. —These robustness tests clusters standard errors at the LLM-level rather than state-level. Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

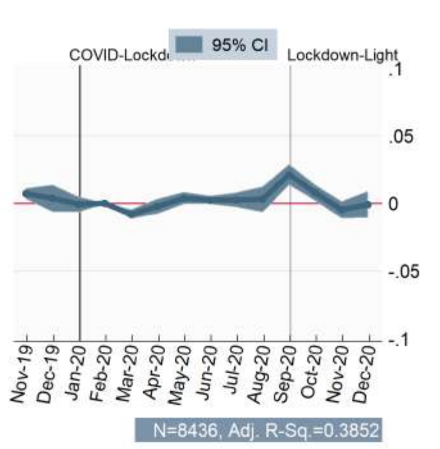
Figure 23: DiD Estimates: LLM-level (*Changing clusters*), 2020-01 - 2020-12



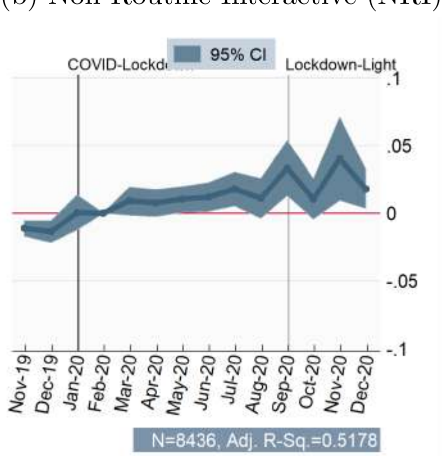
(a) Non-Routine Analytic (NRA)



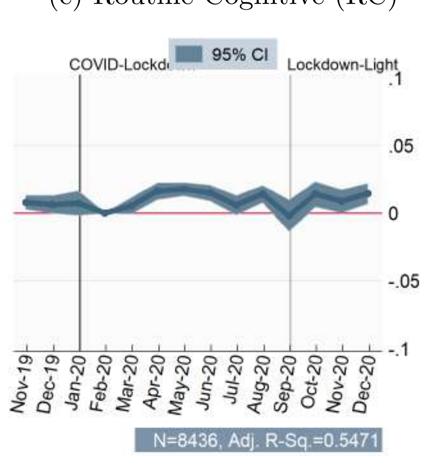
(b) Non-Routine Interactive (NRI)



(c) Routine Cognitive (RC)



(d) Routine Manual (RM)



(e) Non-Routine Manual (NRM)

NOTE. —These robustness tests extend pretrends until November 2019 rather than January 2020. Each coefficient is an estimate of the difference between the YoY growth rate of the local task demand between 2020 and 2019 of the respective month and a weighted geometric average of the YoY growth rates of the 2 previous years, i.e., 2018 and 2017. Coefficients are estimated based on (2). All specifications contain a full set of controls, comprising local skill composition, age composition, citizenship composition, gender composition, industry composition, unemployment rate, and technology differences via WFH suitability.

Figure 24: DiD Estimates: LLM-level (*Extended Pretrends*), 2020-01 - 2020-12