Prague University of Economics and Business

Faculty of Economics

Major specialisation: Economic Analysis

The Impact of Working from Home on Employee Performance during the COVID-19 Epidemic

Diploma Thesis

Author: Iryna Didinova
Supervisor: Ing. Tomáš Miklánek, M.A., Ph.D.
Year: 2021

I hereby declare on my word of honour that I have written the diploma thesis independently, using the listed literature.

Iryna Didinova
In Prague, date 14.05.2021

Acknowledgements

First and foremost I am extremely grateful to my supervisor, Ing. Tomáš Miklánek, M.A., Ph.D. for his invaluable advice, continuous support, and patience during my work on the master thesis. I would also like to thank Ing. Tomáš Lichard, M.A., Ph.D. and Mgr. Milan Bašta, Ph.D. for their support on R software, statistics and econometrics. Finally, I would like to express my gratitude to the Human Resource department of the Firm for providing me the data set for my analysis.

Prague University of Economics and Business
Department of Economics

Faculty of Economics
Academic year: 2020/2021

DIPLOMA THESIS TOPIC

Author of thesis: **Bc. Iryna Didinova**

Study programme: Economics and Economic Administration

Field of study: Economic Analysis

Topic: **The impact of working from home on employee performance during COVID-19 epidemic**

Guides to writing a thesis:

1. 1. This thesis aims to analyze the impact of Stay-at-Home Policies on Work Outcomes of insurance sector employees during the COVID-19 epidemic in Czech Republic using the Instrumental Variable Regression as well as Difference-in-Difference method.
2. 2. COVID-19 has recently become a deadly global issue. Governments around the world are forced to take action to slow the spread of the virus. One of the policies is social distancing. As a consequence of this policy, schools, workplaces, tourist attractions, and even some public transportation have been closed indefinitely. Many companies have followed government regulation to work from home. However, until now the effect of work-from-home on job performance of employees remains debatable thus creating a research gap, hence it is necessary to conduct further research on this undeniably important topic. This diploma thesis aims to analyze the impacts of remote working on performance of middle office and back-office employees within insurance industry with a quantitative approach. Internal anonymized individual level data of real insurance firm will be used. The result of such analysis is valid for at least narrowly defined type of employee: middle office and back office of insurance & banking industry. Moreover, it should provide an useful insight even for many similar types of administrative positions which are very often switched to the home-office environment during the COVID-19 pandemic.
3. In the theoretical part a research of literature connected to remote working and employee productivity will be conducted. While some of the researchers, such Bloom (2015), Rupietta (2017) claim, that working from home increases the productivity of the employees due to higher work satisfaction and lower job attrition rates, other researchers, such as Monteiro et al. (2019) claim the opposite. The results of older pre-COVID articles will be compared with those from 2020 in order to capture the modern trends. Also, the methodology for Instrumental Variables regression as well as Difference-in-Difference method will be introduced in this part of the diploma thesis.
4. In the practical part, the impact of Home Office on productivity of certain firm's employees will be analyzed using the Instrumental Variable approach as well as DIF-in-DIF method. This method compares the difference in outcomes of treated and untreated groups, before and after the Stay-at-Home Policies were introduced in Czech Republic. Employees with no possibility to perform their work remotely will serve as untreated or control group. Percentage fulfillment of individual yearly KPI's will be used as a proxy for employee performance. Instrumental Variable regression will allow to identify the hidden (unobserved) correlation and see the true correlation between the explanatory variable (WFH) and response variable (Performance). Length of person's commuting distance as well as number of children in the household will be used as the Instrumental Variables. These 2 methods will allow to identify the causal effect of working from home on performance.

Length of thesis: 60 stran

Selected bibliography:

1. Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2014). Does Working from Home Work? Evidence from a Chinese Experiment*. The Quarterly Journal of Economics, 130(1), 165–218.
2. Dutcher, E. G. (2012). The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks. Journal of Economic Behavior & Organization, 84(1), 355-363.
3. Monteiro, NP, Straume, OR and Valente, M (2019), 'Does remote work improve or impair firm labour productivity? Longitudinal evidence from Portugal', NIPE Working Papers 14/2019, NIPE – Universidade do Minho.
4. Rupietta, K., & Beckmann, M. (2017). Working from Home. Schmalenbach Business Review, 70(1), 25-55.
5. Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. Econometrica, 65(3), 557.

Diploma thesis topic submission date: February 2021

Deadline for submission of Diploma thesis:May 2021

Bc. Iryna Didinova
Solver

Ing. Tomáš Miklánek, M.A., Ph.D.
Thesis supervisor

prof. Ing. Robert Holman, CSc.
Head of department

prof. Ing. Zdeněk Chytil, CSc.
Dean NF VŠE

Abstract

The novel coronavirus pandemic, currently rushing around the world, has forced governments to impose movement restrictions on their citizens. Authorities implemented local or national lock-downs, work-place risk controls, and business closures. Many companies have followed government regulation to work from home. Given this new situation, organisations across the world are now seeking "the new normal" in which remote working is the standard. In this instance, research on revealed employee performance while working from home is currently more scarce than ever. This study analyzes how working from home impacts employee performance. Data used in the study are internal individual-level anonymized data about employees of real Prague insurance firm in years 2018-2020. The model was fitted using the 2SLS estimator due to the possible endogeneity of Home office participation, caused by self-selection. As a result of such analysis, no impact of working from home on employee performance was found. This result is valid for narrowly defined type of employee: middle office and back office of insurance and banking industry. Moreover, it provides an useful insight even for many similar types of administrative positions.

Keywords: COVID-19, Employee Performance, Employee Productivity, Working from Home, Remote Work, Telework, Telecommuting
JEL Classification: J81, M50, M54

Abstrakt

Nová pandemie koronaviru, která se v současnosti šíří po celém světě, donutila vlády zavést opatření pro omezení pohybu osob. Úřady zavedly místní nebo národní výluky, kontroly rizik na pracovištích a uzavírání podniků. Mnoho společností bylo kvůli vládním nařízením donuceno zavést režim práce z domova. Vzhledem k této situaci nyní organizace po celém světě hledají „nový normál“, ve kterém je vzdálená práce standardem. V tomto případě je výzkum výkonu zaměstnanců při práci z domova vzácnější než kdy dříve. Tato studie analyzuje, jak práce z domova ovlivňuje výkon zaměstnanců. Data použitá ve studii jsou interní, anonymizovaná data na úrovni jednotlivců o zaměstnancích skutečné pražské pojišťovny v letech 2018-2020. Model byl odhadnut pomocí metody 2SLS kvůli možné endogenitě proměnné způsobené tzv "self-selection bias". Na základě této analýzy nebyl zjištěn žádný dopad práce z domova na výkon zaměstnanců. Tento výsledek platí pro úzce definovaný typ zaměstnance: middle office a back office pojišťovacího bankovnictví. Navíc poskytuje užitečný přehled i pro mnoho podobných typů administrativních pozic.

Klíčová skova: COVID-19, Výkon zaměstnanců, Produktivita zaměstnanců, Práce z domova
JEL Klasifikace: J81, M50, M54

Contents

Introduction

The novel coronavirus pandemic currently rushing around the world, has challenged society in ways once considered unimaginable, forcing people to reconsider a wide variety of practices, from work, to leisure, to basic travel and daily tasks. Not only has it changed an individual day-to-day life, but it has also impacted world economy as a whole.

The COVID-19 pandemic has forced governments worldwide to impose movement restrictions on their citizens. Authorities implemented local or national lock-downs, workplace risk controls, and business closures. Many companies have followed government regulation to work from home. Although critical to reducing the virus' reproduction rate, these restrictions come with far-reaching social and economic consequences.

Given this new situation, societies, organisations and workplaces not only in Czech Republic, but across the whole world are now seeking "the new normal" (including the "future of work" and the "role of the office") in which remote working is the norm and people can continue working from home (WFH) in pure or hybrid forms post-COVID-19 (JRC, 2020). Given the potential for cost efficiencies in comparison to the traditional office and the aim of providing employees with more flexibility in choosing where to work, organisations have announced their aim to reduce their office space and introduce blended home–office working conditions post-COVID-19. Insight into how people experience WFH and moreover, how Home office impacts employee's productivity and performance is now even more important than before.

But unfortunately, there is relatively small number of quality studies on working from home and on its impact on performance particularly. Especially scarce is research conducted during the ongoing COVID-19 pandemic. Majority of literature connected to this topic are either exploratory and qualitative studies (Fana et al., 2020; Green et al., 2020; Purwanto et al., 2020; Gajendran and Harrison, 2007a) or studies based on self-reported performance (Ipsen et al., 2021; Chang et al., 2020; Etheridge et al., 2020). Despite the fact, that the later have its advantages, unfortunately the results of such self-reported evidence could often be biased.

Current literature provides little experimental evidence or evidence based on panel data, especially studies written in the age of COVID-19 pandemic. And the rest of articles connected to analysis of remote work impact on employee performance use self-report questionnaires. Thereby, this master thesis is considered to be important contribution to the WFH literature.

First of all, data used in the study are internal individual-level anonymized data about employees of real Prague insurance firm in years 2018-2020, which may be treated as unbalanced panel (some people are the same in all 3 years and some are different, for example employee could have left the Firm or new employees came etc.). The data set collected is considered to be quite big and have in total 1194 observations, in particular 311 in year 2018, 430 in 2019 and 453 in 2020, together with correctly specified model that will allow to receive efficient and unbiased estimate of Home office impact on employee performance.

Possible endogeneity of Home office participation, caused by self-selection, will be treated using Instrumental Variables (IV) estimation strategy. The result of such analysis is valid for at least narrowly defined type of employee: middle office and back office of insurance banking industry. Moreover, it should provide an useful insight even for many similar types of administrative positions which are very often switched to the home-office environment during the COVID-19 pandemic.

In regard to the current literature research, until now the effect of work-from-home on job performance of employees remains debatable. While some of the researchers, such Bloom et al. (2013), Rupietta and Beckmann (2017) claim, that working from home increases the productivity of the employees due to higher work satisfaction and lower job attrition rates, other researchers, such as Monteiro et al. (2019) claim the opposite. On the other hand, some authors suggest, that the impact of remote working on employee performance or productivity can be different in the different situations (Dutcher, 2012; Bao et al., 2020). For example Dutcher (2012) discovers that the remote working environmental effects may have positive implications on productivity of creative tasks but negative implications on productivity of dull tasks. Based on the current research, my hypothesis is that working from home is on average positively associated with performance of the insurance Firm employees.

This thesis is divided into 2 parts. The Theoretical part contains the fist 3 chapters and the Practical part have further two. In the fist chapter of this thesis main information about COVID-19 pandemic is presented. The fist sections provides the main facts about the virus. The second one narrates about the spread of the virus in the Czech Republic as well as about anti-virus measures set up by Czech government during the year 2020 and 2021. The last section summarises the main implications of the pandemic on the prevalence of working from home in Europe.

In the second chapter research of the literature connected to remote working wis conducted. Section one describes theoretical background of employee intrinsic motivation and work performance, presenting the Oldham and Hackman's job characteristics model (1976). The second section of this chapter is dedicated to empirical evidence in regard to the impact of remote working on job performance and is divided to four parts.

The first part summarizes advantages and disadvantages of working remotely, the second part is focused on experimental or panel data studies connected to the field, the third subsection is dedicated to latest findings on working from home during the COVID-19 pandemic and finally the last one describes the contribution of this master thesis to the literature.

For the purpose of this master thesis Instrumental variables regression estimated by the two-stage least squares (2SLS) estimator will be used due to possible endogeneity in the model. Hence methodology for Instrumental Variables estimation is introduced in the third chapter of the diploma thesis. The first section explains , what is endogeneity problem in regression analysis. In the second section econometric theory of Instrumental Variables regression is presented, and finally some interesting examples of instrumental variable use from economic literature are introduced.

In the Practical part of the master thesis I describe the data set used for this study as well as estimate the model and comment the results. In the 4th chapter of this master thesis the data set used for analysis of employee performance is presented. The first section describes, how the data were acquired. In the second section main information about all variables as well as justification for their use is provided. And the third section of the chapter presents description statistics of data set together with graphical representation and further details about data and variables.

Finally in the last chapter the results of analysis are presented. The fists section describes the empirical strategy, the second one provides justification for instrumental variable use as well as presents the results of the First stage regression. The next two sections are dedicated to discussion of the results and also their possible limitations. Finally last section of this chapter explains, why Difference-in-Difference estimator was not used for analysis of employee performance.

Part I

Theoretical part

Chapter 1

COVID-19 pandemic and remote working

In this chapter main information about COVID-19 pandemic is presented. The fist sections provides the main facts about the virus. The second one narrates about the spread of the virus in the Czech Republic as well as about anti-virus measures set up by Czech government during the year 2020 and 2021. The last section summarises the main implications of the pandemic on the prevalence of working from home in Europe.

1.1 COVID-19 worldwide

The novel coronavirus or COVID-19 pandemic, currently rushing around the world, has challenged society in ways once considered unimaginable, forcing people to reconsider a wide variety of practices, from work, to leisure, to basic travel and daily tasks. Not only has it changed an individual day-to-day life, but it has also impacted world economy as a whole.

Coronavirus disease also known as COVID-19 is a contagious disease caused by severe acute respiratory syndrome SARS-CoV-2. It was first identified in Chinese city Wuhan in December 2019. The disease has since spread worldwide, leading to an ongoing pandemic. As of 7 May 2021, more than 156 million cases have been confirmed, with more than 3.25 (“ArcGIS”, n.d.) million deaths attributed to COVID-19, making it one of the deadliest pandemics in history.The World Health Organization declared a pandemic on 11 March 2020 (“WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020”, March 2020).

Currently, the known main mode of transmission is through respiratory droplets, and hence the virus is considered to spread through close contact with other people. World Health Organization recommended various measures in order to prevent further COVID-19 spread in the population. Preventive measures include social distancing and wearing face masks as well as covering one’s mouth when sneezing or coughing, frequent hand disinfecting or washing, persistent premises ventilation and air-filtering, disinfecting surfaces etc. Furthermore self-isolation and monitoring for symptomatic people or people exposed to the COVID-19 virus is recommended (“Advice for the public on COVID-19”, March 2020).

Several COVID-19 vaccines have been already developed and are now widely distributed. Unfortunately, no cure, which is able to inhibit the virus, has been developed yet, hence current treatments focus only on addressing COVID-19 symptoms (Rogers, March 2020).

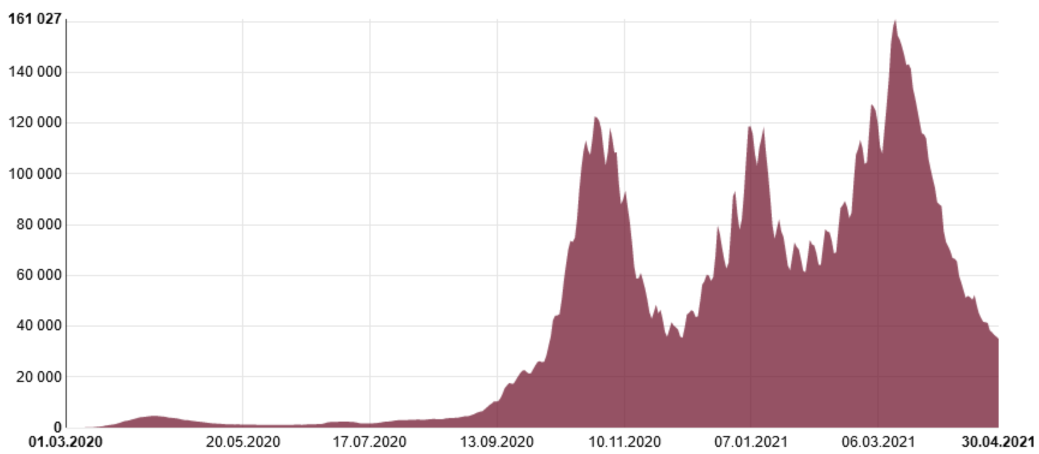
The COVID-19 pandemic has forced governments worldwide to impose movement restrictions on their citizens. Authorities implemented local or national lock-downs, workplace risk controls, and business closures. Many companies have followed government regulation to work from home. Although critical to reducing the virus' reproduction rate, these restrictions come with far-reaching social and economic consequences.

1.2 COVID-19 in Czech republic

COVID-19 was firstly confirmed in Czech Republic on 1 March 2020 ("V České republice jsou první tři potvrzené případy nákazy koronavirem", March 2020). Northern Italy was stated to be the source of the decease, since it used to be a popular resort amongst Czech residents. On 12 March, for the first time in the Czech Republic's modern history, the government declared a nationwide state of emergency ("Measures adopted by the Czech Government against the coronavirus", 2021). Soon after, the state closed its borders and issued a nationwide curfew. Wearing face masks in public become mandatory for all citizens (Czech Republic was the first European country to introduce this measure ("Could Czech's Measure to Fight Coronavirus Save Thousands of Lives?", 2020)). All establishments, except for the most essential ones, like supermarkets and pharmacies, were closed for customers.

A general closure of services and retail sale was in place until May 11th ("Measures adopted by the Czech Government against the coronavirus", 2021); however, all shops could conduct socially-distances sales with delivery through improvised temporary take-out windows. Majority of companies strictly recommended employees to utilize work from home as much as possible or even forbid coming to office. Universities, schools and kindergartens were closed as well. The nationwide state of emergency in Czech republic ended on 17 of May 2020 ("Measures adopted by the Czech Government against the coronavirus", 2021), overall it lasted 66 days.

Figure 1.1: Active COVID-19 cases in Czech Republic (by 30.04.2021)



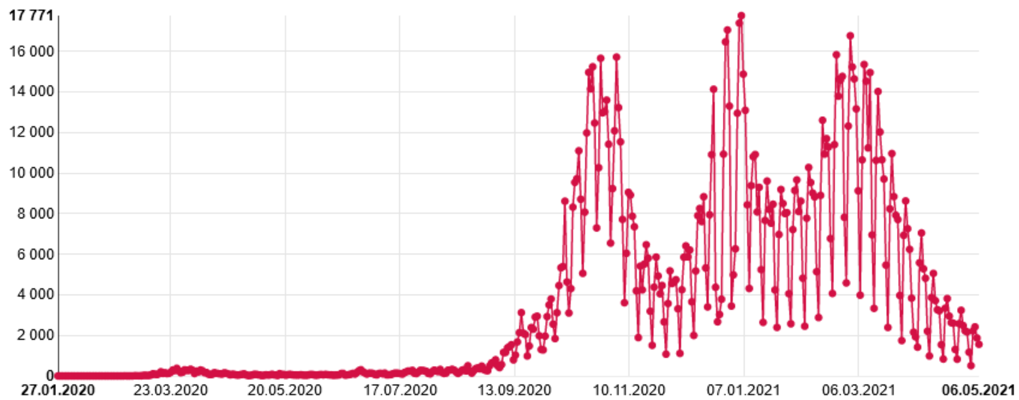
Source:"COVID-19: Onemocnění Aktuálně MZČR", n.d.

During the summer 2020 almost no pandemic measures were applied in Czech Republic, as a result the number of infected by COVID-19 as well as number of COVID-19 attributed deaths was growing very quickly. As can be seen from figure (1.1), the number of people infected during the second wave of the pandemic was multiple times higher than during the first one. The development of daily new cases (see figure 1.2) and deaths has more or less similar trend as the development of active cases.

As a result of quick decease spread, the government declared a nationwide state of emergency for the second time, and it lasted from 5 October till 11 April (almost half a year!). Moreover, a new, more transmittable variant of COVID-19, Lineage B.1.1.7, emerged in the Czech Republic in January 2021 ("UPDATE 1-Czech Republic detects UK coronavirus variant, to maintain lockdown measures", 2021). As a result the 3rd and 4th pandemic waves in Czech Republic followed quickly one by one. Very high increases on daily number of infected took place during this period (which was achieved partially due to very intensive COVID-19 testing), often it reached values even higher than 10 000. Daily new COVID-19 cases reached its peak with value 17 771 on 6 January ("COVID-19: Onemocnění Aktuálně MZČR", n.d.).

More or less strict movement restrictions were imposed on Czech citizens during the whole time of the second national state of emergency. Pandemic measures were overall similar to those imposed during the first wave: closed services, retail sale, restaurants, universities, borders closed for foreigners (with exceptions), nationwide curfew etc. Unlike during the first wave of COVID-19, kindergartens were open for majority of time, elementary schools were sometimes opened for the first stage pupils. The Czech government also prohibited unessential travelling between districts as of 1 March ("Ministerstvo vnitra České republiky", 2021).

Figure 1.2: daily new COVID-19 cases in Czech Republic (by 30.04.2021)



Source: "COVID-19: Onemocnění Aktuálně MZČR", n.d.

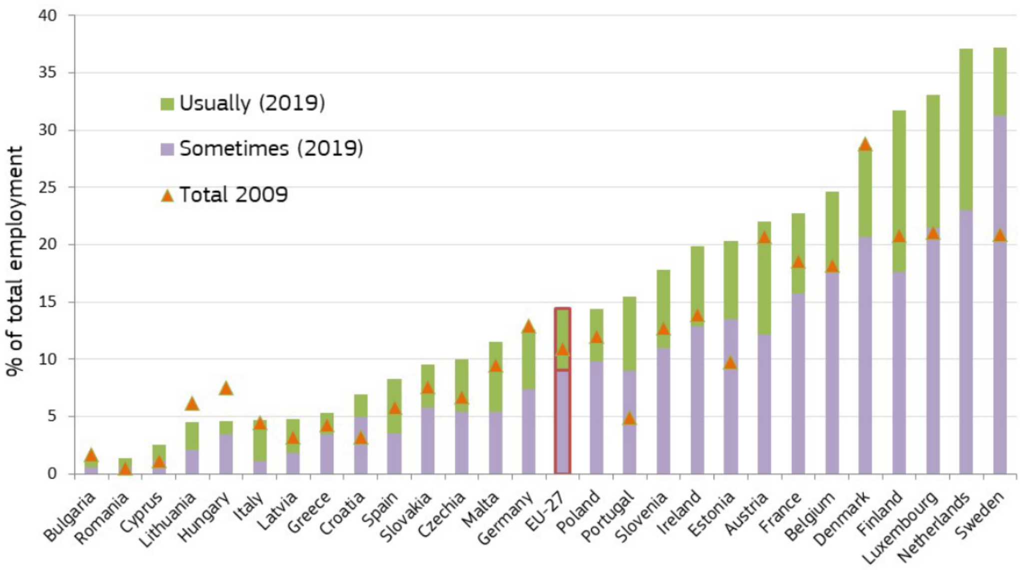
1.3 Prevalence of remote working before and after the COVID-19 pandemic

It is appropriate to provide a definition of the term ”remote working” in the beginning of this section. According to the Cambridge Dictionary:

”Remote working is the practice of an employee working at their home, or in some other place that is not an organization's usual place of business

Unfortunately, there is no generally accepted definition for remote working in the literature yet and it has been studied under various names, in particular: ”teleworking”, “telecommuting”, ”home office”, ”working from home” etc. All these terms are currently used interchangeably in the literature. In this study terms ”remote work” and ”working from home (WFH)” will be used as they are considered to be the most appropriate for the purpose od the thesis.

Figure 1.3: Prevalence of remote working across EU Member States in years 2009 and 2019



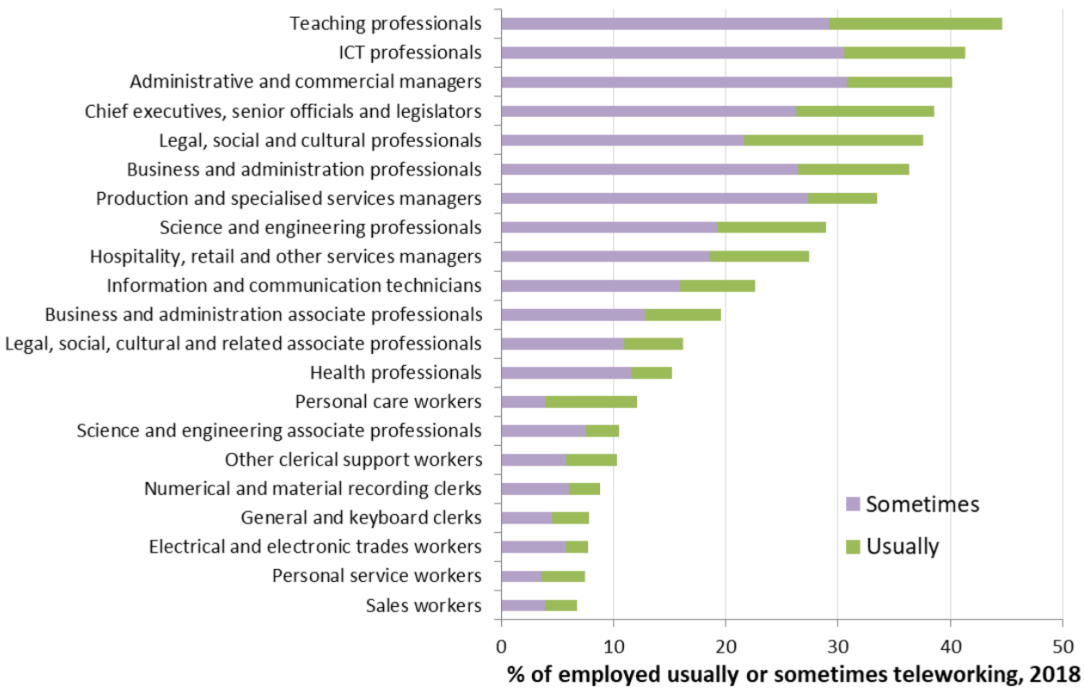
Source: JRC, 2020 April-May

Before the pandemic, discussions on the future of work-life were unclear and often questioned, but COVID-19 forced a decision upon people. In order to lower the risk of spreading the virus, many firms were encouraging it's employees to work from home as much as possible. According to Eurostat's COVID-19 survey (2020), during the first wave of epidemic in Czech republic one third of all employees (34.1% exactly) was working from home (see figure (1.5)), during the summer this number was even higher, and was equal to 44.3%. The number of people using Home office increased radically since the beginning of COVID-19 pandemic.

Figure (1.3) shows, that around 10% of total employment in Czechia (JRC, 2020) was working from home usually or at least sometimes, which is slightly lower, than EU-27 level.

As can be seen from the figure (1.3), the European countries are quite heterogeneous in prevalence of remote working in 2019, it varies from close 0% in Romania and Bulgaria and 37% in Netherlands and Sweden. In regard to abundance of WFH between different occupations in EU-27, it can be seen from figure (1.4), that WFH is most commonly used by teaching and ICT (Information and communications technology) professionals as well as administrative and commercial managers - 40% (!) of employed in this occupations work from home usually or at least sometimes. Unfortunately, the very nature of some occupations makes it difficult or impossible to perform them away from the standard worksite. This is generally the case of activities that involve a high level of face-to-face interaction with the public, for example sales workers, servers, or personal service workers such as hair stylists, who showed before the pandemic the lowest shares of telework among major occupational groups.

Figure 1.4: Prevalence of remote working by occupation in 2018, EU-27



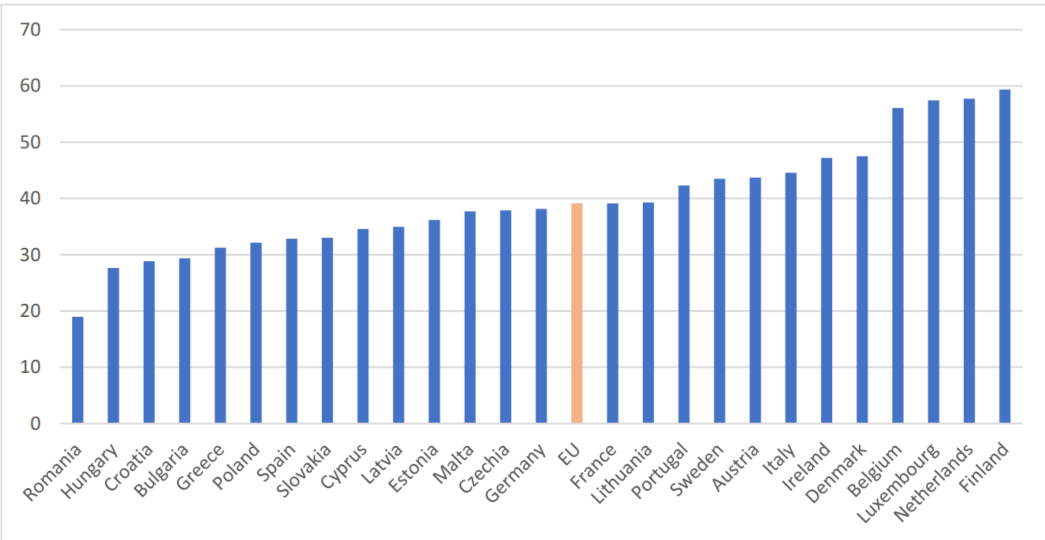
Source: Sostero, 2020

Prevalence of remote working due to COVID-19 have increased not only in Czech Republic but an all European Union. According to Eurostat, as of 2019, only 5.4% of employed in the EU-27 usually worked from home – a share that remained rather constant since 2009. However, over the same period, the share of employed working at least sometimes from their homes increased from 5.2% in 2009 to 9% in 2019 (JRC, 2020). But the situation changed drastically after the outbreak of the Covid-19 pandemic, and working from home has become the norm for millions of workers in the EU and worldwide.

Early estimates from April 2020 Eurofound's COVID-19 survey (Eurostat, 2020;Sostero, 2020) suggest that close to 40% (see figure (1.5)) of those currently working in the EU began to telework fulltime as a result of the pandemic. Later JRC study (Eurostat, 2020) provides a rough estimation of around 25% of employment in teleworkable sectors in the EU as a whole.

Given this new situation, societies, organisations and workplaces not only in Czech Republic, but across the whole world are now seeking "the new normal" (including the "future of work" and the "role of the office") in which teleworking is the norm and people can continue working from home (WFH) in pure or hybrid forms post-COVID-19 (JRC, 2020). Given the potential for cost efficiencies in comparison to the traditional office and the aim of providing employees with more flexibility in choosing where to work, organisations have announced their aim to reduce their office space and introduce blended home–office working conditions post-COVID-19. Insight into how people experience WFH and moreover, how Home office impacts employee's productivity and performance is now even more important than before.

Figure 1.5: Employees working from home during COVID-crisis (April-May) in %



Source: Sostero, 2020

Evidence suggests that in normal times people working from home can sustain, or even enhance, their productivity, while enjoying a better work-life balance. Yet, under the current exceptional circumstances productivity, working conditions, or both, may be deteriorating for many workers due to, among other problems, lack of childcare, unsuitable working spaces and ICT tools (JRC, 2020). In this instance, it is necessary to conduct further research on this undeniably important topic. consequently, this diploma thesis aims to close such research gap at least partially by analyzing the impacts of remote working during COVID-19 epidemic on performance of middle office and back-office employees within insurance industry in Czech republic.

Chapter 2

Literature review

The prevalence of working from home in firms has increased over the past decades due to advancements in communication and information technologies. In recent years, with the widespread use of cloud services and remote access to work applications, employees can easily perform their tasks outside the office. Currently, in times of ongoing COVID-19 pandemic, the majority of firms in Czech Republic enable their employees work at least partially from home, and some even give the possibility to work remotely full-time. Remote workers may work at home but also turn to coffee shops or co-working spaces, or even travel around the world while maintaining their career goals. Video conferencing permits to communicate and interact with colleagues in real time in any place with more or less stable internet connection.

In this chapter research of the literature connected to remote working will be conducted. Section one describes theoretical background of employee intrinsic motivation and work performance, presenting the Oldham and Hackman's job characteristics model (1976). The second section of this chapter is dedicated to empirical evidence in regard to the impact of remote working on job performance and is divided to four parts. The first part summarizes advantages and disadvantages of working remotely, the second part is focused on experimental or panel data studies connected to the field, the third subsection is dedicated to latest findings on working from home during the COVID-19 pandemic and finally the last one describes the contribution of this master thesis to the literature.

2.1 Theoretical background

Recent Technical Report (May 2020) by Joined Research Center uncovered that 37% of dependent employment (16) in the EU is currently possible to perform remotely – very close to the estimates of remote working indicated in real-time surveys during the COVID-19 crisis. Because of differences in the employment structure, the fraction of employment, possible to perform remotely, ranges between 33-44% in all but five EU member states (Sostero et al., 2020). Some jobs, especially those related to healthcare, farming and hospitality cannot be performed at home.

Regardless of increasing popularity, so far no theoretical work has explicitly modelled the linkage between remote work and employee or firm performance. Past empirical research has borrowed various arguments and mechanisms from different strands of economics and related fields to explain the referred linkage.

For example model of Hackman and Oldham, 1976, specifies the conditions under which individuals will become internally motivated to perform effectively on their jobs. The model was tested for 658 employees who work on 62 different jobs in seven organizations, and results support its validity. The model is depicted on a diagram (2.1) below. As can be seen from the figure (2.1), 5 key job dimensions induce 3 psychological states which, in turn, lead to a number of beneficial personal and work outcomes, and this relationships between them are constrained by individual growth need strength. The three psychological states are considered to be the causal core of the model, and they indicate that individual will become internally motivated to perform effectively in case, *that he/she learns (knowledge of results) that he personally (experienced responsibility) has performed well on a task that he cares about (experienced meaningfulness)* (Hackman and Oldham, 1976).

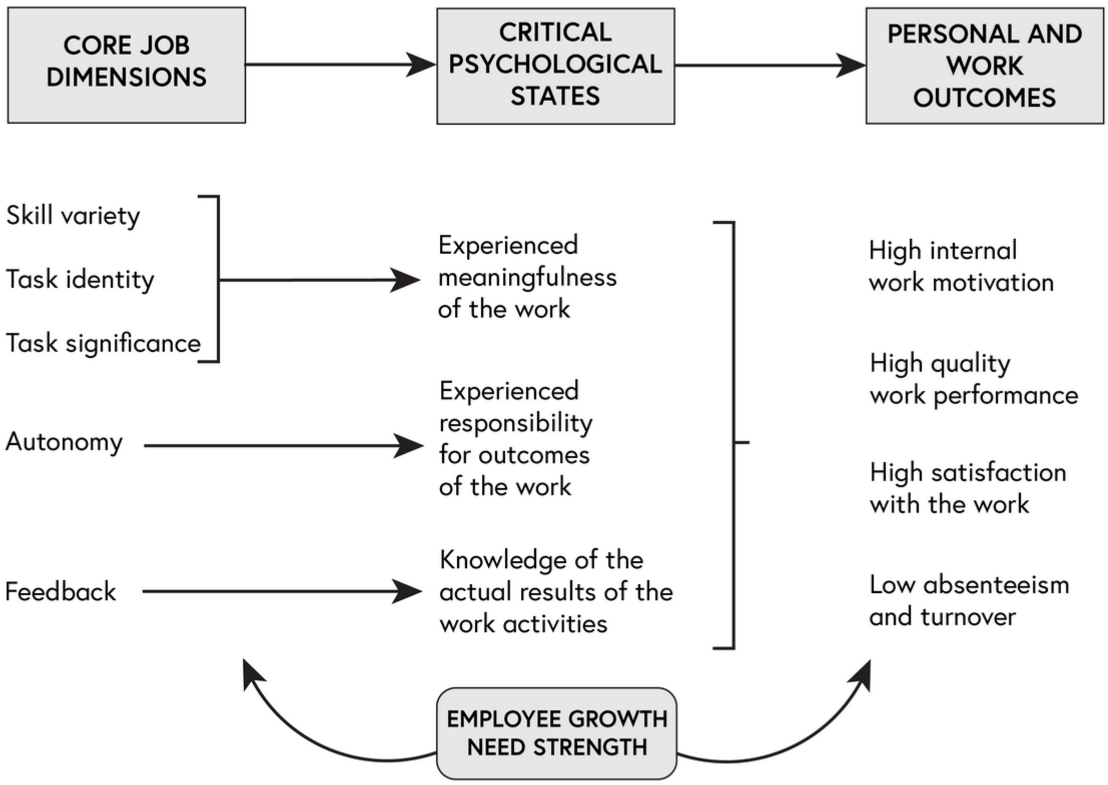
To summarize, the overall potential of a job to induce internal work motivation should be the highest once all of the following are true:

- the job is high on at least one of the 3 job dimensions leading to experienced meaningfulness;
- the job is high on autonomy;
- the job is high on feedback (Hackman & Oldham, 1976).

The authors also created measure of the degree to which the above conditions are met (2.1), which was called The Motivating Potential Score (MPS). As can be seen from the formula, a near-zero score of a job on either autonomy or feedback will reduce the overall MPS to near-zero, while a near-zero value of one of the three job dimensions which promote experienced meaningfulness cannot do so.

$$MPS = \left[\frac{SkillVariety + TaskIdentity + TaskSignificance}{3}\right] \times Autonomy \times Feedback \quad (2.1)$$

Figure 2.1: Oldham and Hackman's job characteristics model



Source: Hackman and Oldham, 1976

In regard to connection between Oldham and Hackman's model and working from home, possibility of remote working increases employees' autonomy in scheduling and organizing their work, as employers have fewer possibilities for control. Hence, remote work is considered to affect intrinsic motivation and consequently work performance positively.

2.2 Impact of remote working on performance: Empirical evidence

2.2.1 Advantages and disadvantages of working remotely

Although the popularity of WFH has increased worldwide, academics argue regarding its pros and cons. According to exploratory study ”Managing Telework Programs Effectively”, written in 2007 by Jeffrey A. Mello, remote working could bring a lot of beneficial effects for employees, for example spared commuting time, better work-life balance, increased motivation, improved gender diversity (e.g. women and careers), better productivity and higher job satisfaction.

Moreover, remote working can be beneficial for employers as well: allowing employees to work from home could reduce office rental costs (less office space needed, in case employees are working from home at least partially). Furthermore WFH can potentially decrease employee attrition rates and absenteeism, increase talent retention as well as it can enhance employment opportunities (for example new talents from around the world) (Mello, 2007).

Additionally, WFH have also certain societal benefit, in particular, WFH could be a way out for individuals with disabilities, who are qualified to hold a myriad of jobs, but unfortunately, have difficulties physically getting to workplace. In addition, remote work can be also suitable for elderly, who may have limited physical mobility, or new parents, who need or prefer to be at home with a child (Mello, 2007).

There is also a connection to the urban economics literature. On top of all above mentioned benefits, remote work could definitely bring one more, in particular - WFH has positive impact on environment. For instance, it enables to reduce traffic and, as a result, decreases levels of pollution by diminishing the number of vehicles on the road. Moreover, widespread remote work in population would reduce strain on public transport systems (Harpaz, 2002) as well as it would reduce population centrality as people move out to the suburbs (Bento et al., 2005).

On the other hand, literature points out quite a few drawbacks of working remotely. WFH could possibly improve work-life balance of employee, but on the other hand there is a risk of blurred lines between work and personal life and potential for work/family conflict. WFH prevents from office distractions or office noise, but instead provides with family distractions. Moreover, WFH is difficult once a person leaves in small place with other people and does not have a proper work place (Vyas and Butakhieo, 2021). Which is quite a big issue in current times, when due to COVID-19 pandemic schools, kindergartens and offices are often closed, and whole family must somehow survive in one flat and in addition perform their tasks efficiently. Furthermore, remote working has further drawbacks for an employee such as lack of hardware support and supervision, communication barriers and lack of proper communication. All above mentioned could decrease employee productivity quite significantly.

There are also major employer-level drawbacks of WFH, in particular cybersecurity (loss of corporate/employee/customer data) (Vyas and Butakhieo, 2021) and potential for lower performance or quality of work without direct supervision of manager or due to lack of proper communication as was mentioned above (for example, employee didn’t understand the task correctly etc.)

2.2.2 Employee performance and remote working

It is important to highlight, that foregoing advantages of working from home were taken from mostly exploratory studies, but is there any hard evidence to support such findings?

One of the most influential works in this field, is a study conducted by Bloom et al., 2013, in Chinese 16,000-employee travel agency named Ctrip. It is considered to be the first randomized field experiment on WFH. The research was conducted due to the fact, that company's senior management was interested in allowing its Shanghai call center employees to work from home to reduce office rental costs, which were increasing rapidly due to the thriving real estate market in Shanghai. Senior management also believed, that introducing WFH could potentially bring down high attrition rates by saving the employees from prolong travelling to workplace every day. But on the other hand, management was afraid, that introduction of WFH will reduce employee's productivity without direct supervision of superordinates.

In summary, Ctrip decided to run a nine-month experiment on WFH. 249 eligible Call center employees were randomly assigned to either treatment or control group. The only difference between the two groups was the location of work: treatment employees were supposed to work 4 days at home and 1 day at the office, while control employees worked 5 days at the office.

The authors findings were prominent: the performance of the treatment group went up significantly - by 13% over 9 months of the experiment, where 9% increase was attributed to overall longer shifts worked (reductions in breaks and seek days), and remaining 4% rise was affixed to increased productivity per 1 minute (higher number of calls per minute). In interviews, the workers attributed the increase in time worked to the greater convenience of being at home (e.g., the ease of getting tea, coffee, or lunch or using the toilet) and the increased output per minute to the relative quiet at home. Moreover, researches found, that attrition rates went 50% down in comparison with the control group. In addition, Bloom did not found any evidence of worsened quality of home working.

Experiment results in regard to individual-level worker productivity were indeed striking. But how did WFH impacted firm-level productivity of Ctrip agency? Thanks to introducing WFH,the firm improved total factor productivity by between 20% to 30% and saved about 2,000 a year per employee WFH. About two thirds of this improvement came from the reduction in office space and the rest from improved employee performance and reduced turnover.

Researchers also found one disadvantage of WFH - 50% reduction in promotion rates conditional on performance. There was also one more interesting finding. After the end of experiment Ctrip introduced a possibility of remote forking to the entire company.

Surprisingly, over halve of employees, who participated in experiment (initially they all volunteered to work from home) decided to stay in the office after the end of experiment, citing concerns over the loneliness of home working. After correction for re-selection, researchers claim, that impact of WFH on employee performance increased from 13% to 22% (Bloom et al., 2013).

Such results could mean, that remote working is not suitable for everybody. Indeed, some employees with smaller need for social interactions could experience better productivity while working from home. But the fact that 2/3 of control and half of the treatment group decided to work in the office after the end of the experiment, could mean that the majority of people is simply not fit for remote working? On the other hand, it is also possible, that some employees would like to WFH, but they did not do so due to 50% decline in promotion rates as a result of WFH. The other explanation may be the fact, that positive effects of WFH are temporary, and employees became tired after 8 months of remote work.

Besides, it is unknown, what effect will remote working have on people with other occupations. After all, call center employees are considered to be particularly suitable for WFH, because their job requires neither teamwork no in-person face time. Quantity and quality of employee performance can be easily quantified and evaluated. The link between effort and performance is direct.

There is an experimental study (Dutcher, 2012), which could shed some light on above mentioned research question. Dutcher examined, how dull and creative tasks influence productivity of people, when they are working remotely. The experiment was conducted on students at Florida State University. In total there were 125 people, which were split to treatment (worked at home) and control group (worked in lab). Every student was supposed to complete 2 task, first one was ”dull” (typing random sets of characters) an second one was ”creative” (coming up with unusual uses for common objects) [1] Results of the study suggests that the remote working environmental effects may have positive implications on productivity of creative tasks but negative implications on productivity of dull tasks. However, due to the experimental nature this findings have limited validity for the population.

One more question pops out in regard to results of Bloom’s experiment: does the frequency of WFH matter? Rupietta and Beckman in their study ”Working from Home – What is the Effect on Employees’ Effort” (2016) found out, that the more often employees work from home, the higher is the work effort they provide. The work was based on the data from the German Socio-Economic Panel (SOEP) and Instrumental variables estimation strategy was used in order to account for self-selection into working locations. Employee effort was measured as the difference between average actual working hours and contractual working hours for employee. The empirical results of the given study show, that working from home has a significantly positive influence on work effort (Rupietta and Beckmann, 2017). But due to the very nature of dependent variable, it is arguable, that the results are indeed valid, because it is possible, that difference between actual and contractual working hours isn’t given by increased working effort, but due to lower productivity during WFH, worse time managements and distractions.

[1] This task is known as “unusual uses test” and it has been used in psychology since 1962 (Torrance, 1962) to measure divergent thinking. In this test subjects saw a word or phrase describing an object along with space to type a use for which the object was not intended.

Even though quite a few studies have found positive impact of remote working on employee performance or productivity, Monteiro et al. (2019) , claims the opposite. The authors suggest, that the reason of such positive results may lie in data set used for analysis, which is often cross-sectional data from single firm or industry. Costrastingly, Monteiro et al. (2019), used rich longitudinal panel data set of Portuguese companies, collected over the period 2011-2016. The data set included representative sample of the whole economy, including manufacturing and service industries.

The empirical strategy was based on the augmented Cobb-Douglas production function (Bloom et al., 2019 as cited in Monteiro et al., 2019).

$$\ln\left(\frac{Y}{L}\right)_{it} = \alpha \ln\left(\frac{K}{L}\right)_{it} + \beta \ln\left(\frac{M}{L}\right)_{it} + \gamma \ln L_{it} + \theta RW_{it} + \delta' Z_{it} + \upsilon_t + \epsilon_{it}, \quad (2.2)$$

where Y_{it} is the value of output measured by total revenues/sales, K_{it} is the value of total tangible assets, M_{it} are intermediate inputs, and L_{it} is the number of workers in a company i at time t. In addition, RW_{it} is an indicator variable, which points out weather at time t a company i enables remote work, Z_{it} is a vector of observable control variables, υ_t controls for time-specific shocks common to all firms, and finally ϵ_{it} is an error term.

Finally, the authors use difference-in-difference estimator (see section 5.5) based on firmed fixed effects in order to estimate the impact of remote working on productivity, in particular treated companies (firms which either adopt or abandon remote working during the period of observation) were compared to untreated companies (which never or always allow remote work). Moreover, the authors also include industry specific time trends in the equation to allow for different technological progress by industry and to control for industry-specific business cycle effects.

Monteiro et al. (2019), discovered statistically significant but small negative effect of enabling remote work in firms. Nevertheless, they also found substantial heterogeneity across different firm categories, suggesting that companies undertaking Research and Development (RD) activities have nonetheless positive effect on labour productivity and otherwise. The authors also discovered that small firms, which do not export and have below-average skill level workforce, are also prone to negative effects of remote working.

There is one more interesting study (Bao et al., 2020), which uses quantitative approach in order to find out, how does working from home effects developer productivity during COVID-19 pandemic in China. The authors used the records about 139 developers' activity during the 138 working days, part of which was collected during the COVID-19 lockdown. The researches discovered both negative and positive impacts of WFH on developer productivity while using different metrics (the number of builds/commits/code reviews). The authors also found out, that WFH has different impacts on projects with different characteristics including programming language, project type/age/size. For example, WFH might have a positive effect on developer productivity for small project but the effect was negative for large projects.

2.2.3 Remote working during the pandemic: Research on self-reported performance

For sure the impacts of WFH on employee performance could be different during the times COVID-19 pandemic. In this instance, the paper from Centre for Economic policy research (October 2020) could provide some interesting insites on home workers' self-reported productivity towards the end of the lockdown period in the UK (June 2020). The data used for analysis were individual-level and were taken from the UK Household Longitudinal Survey (UKHLS). The study discovered that on average workers report being approximately as productive as before the pandemic. However, productivity varies substantially across socioeconomic groups, industries and occupations (Etheridge et al., 2020).

In particular, employees in industries and occupations, which are less suitable for remote working, as well as women, low earners and self-employed reported lower productivity than before the pandemic and vice versa. Their low productivity is not only related to their job characteristics, but is also directly affected by their socioeconomic conditions (Etheridge et al., 2020). For instance, according to Adams-Prassl et al. (2020), women are not only more likely to work in occupations less suitable for home work, but also their productivity is more negatively affected by the presence of children. Feng and Savani (2020), also argues, that COVID-19 lockdown created gender gap in perceived work productivity and job satisfaction. The authors used a sample of employed women and men from dual career families from United States, who had at least 1 child and were working from home since Covid-19 pandemic emerged. It was found that before the Covid-19 pandemic, there were no gender differences in self-reported work productivity and job satisfaction. However, during the lockdown, women reported lower work productivity and job satisfaction than men. The authors claim, that the reason for such results was the fact, that childcare was not available during the lockdown, and women were expected to devote more time to housework and childcare.

Etheridge, Tang and Wang found as well that those who previously worked from home at least sometimes and then increased the intensity of home-working experienced a productivity increase. On the other hand, those who did not increase their home-working frequency or never worked from home before the pandemic report a large productivity decline (Etheridge et al., 2020). Such evidence suggests the presence of 2 opposite forces: a positive productivity effect of increased remote working alongside a direct negative effect of the pandemic itself. The productivity decline reported by those who have always worked from home is evidence of such phenomenon.

Another study on self-reported performance (Ipsen et al., 2021), which was based on large data set collected from European knowledge workers' during the early stages of lockdown (11 March to 8 May 2020), also discovered on average positive experience with WFH. According to the authors, the main advantages were improved work–life balance, higher work efficiency and greater work control, and the main drawbacks were home office constraints, work uncertainties and inadequate tools.

Unfortunately, all above mentioned studies were conducted on data from early stage of COVID-19 pandemic, and the results of prolonged working from home are still unknown.

2.2.4 Contribution to the literature

Unfortunately, there is relatively small number of quality studies on working from home and on its impact on performance particularly. Especially scarce is research conducted during the ongoing COVID-19 pandemic. Majority of literature connected to this topic are either exploratory and qualitative studies (Fana et al., 2020; Green et al., 2020; Purwanto et al., 2020; Gajendran and Harrison, 2007a) or studies based on self-reported performance (Ipsen et al., 2021; Chang et al., 2020; Etheridge et al., 2020). Despite the fact, that the later have its advantages, unfortunately the results of such self-reported evidence could often be biased.

One of the main advantages of the self-report questionnaire is that it can be administered to a large sample of people quickly without much effort or financial cost. Probably promptness and simplicity of such research is the reason, why there are already so many questionnaire studies on WFH during COVID-19 and no articles based on revealed evidence, data for which are more difficult to collect. Generalization of the questionnaire findings is possible when a quantitative data sample is big and especially when it is collected randomly (Demetriou et al., 2015).

Nevertheless the main disadvantage of self-report questionnaires might be the possibility of providing invalid answers. There are various biases which could occur (social desirability bias, response and non-response bias etc.), in addition a questionnaire itself could be design incorrectly, which may lead to obtaining different interpretations of questions or the structure might force participants to answer in a way that does not match their views (Demetriou et al., 2015; Rosenman et al., 2011).

Current literature provides little experimental evidence or evidence based on panel data, especially studies written in the age of COVID-19 pandemic. And the rest of articles connected to analysis of remote work impact on employee performance use self-report questionnaires. Thereby, this master thesis is considered to be important contribution to the WFH literature. First of all, data used in the study are internal individual-level anonymized data about employees of real Prague insurance firm in years 2018-2020, which may be treated as unbalanced panel (some people are the same in all 3 years and some are different, for example employee could have left the Firm or new employees came etc.). The data set collected is considered to be quite big and have in total 1194 observations, in particular 311 in year 2018, 430 in 2019 and 453 in 2020, together with correctly specified model that will allow to receive efficient and unbiased estimate of Home office impact on employee performance.

Possible endogeneity of Home office participation, caused by self-selection, will be treated using Instrumental Variables (IV) estimation strategy. The result of such analysis is valid for at least narrowly defined type of employee: middle office and back office of insurance banking industry. Moreover, it should provide an useful insight even for many similar types of administrative positions which are very often switched to the home-office environment during the COVID-19 pandemic.

COVID-19 have shaped the world and remote working is going to stay with us at least in the form of hybrid working patterns. Companies are eager to improve cost-efficiency via reduction in office space, while providing employees with more flexibility in choosing where to work. In this instance, research on revealed employee performance while working from home is currently more scarce than ever.

Chapter 3

Instrumental Variables approach

For the purpose of this master thesis Instrumental variables regression estimated by the two-stage least squares (2SLS) estimator will be used, as variable representing number of days spent working from home in particular year is expected to be an endogenous variable. Hence methodology for Instrumental Variables estimation is introduced in this chapter of the diploma thesis. The first section explains , what is endogeneity problem in regression analysis. In the second section econometric theory of Instrumental Variables regression is presented, and finally some interesting examples of instrumental variable use from economic literature are introduced.

3.1 Endogeneity problem

In traditional economic usage, a variable is endogenous if it is determined within the context of a model, but words "endogeneous variable" in econometrics have somewhat different meaning: it is used broadly to describe any situation where an explanatory variable is correlated with the disturbance (J. Wooldridge, 2002). The concept can be explained more in depth using ordinary least squares estimation (OLS) as an example.

Let's consider a linear in parameter's population model

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + u \tag{3.1}$$

where $y, x_1, x_2, ..., x_k$ are observable random scalars, $\beta_1, \beta_2, ..., \beta_k$ are parameters to be estimated and u is unobservable random error.

In order for OLS to provide consistent estimates of β_j the population error term denoted as u has to have zero mean and be uncorrelated with each of the regressors:

$$E(u) = 0, COV(x_j, u) = 0, j = 1, 2, ..., K \tag{3.2}$$

Sufficient for assumption (3.2) is the zero conditional mean assumption

$$E(u|x_1, x_2, ..., x_K) = E(u|\mathbf{x}) = 0 \tag{3.3}$$

Under equation (3.1) and assumption (3.3) there is the population regression function (J. Wooldridge, 2002):

$$E(u|x_1, x_2, ..., x_K) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k \tag{3.4}$$

In the case, when assumption (3.3) does not hold, endogeneity problem emerges, as a result OLS fails to provide consistent estimates of all the β_j, then regression estimates measure only the magnitude of association, rather than the magnitude and direction of causation, both of which are needed for analysis. When OLS is inconsistent due to endogeneity of its regressor(s), the changes in x_j are associated not only with changes in y but also in error term u (Cameron and Trivedi, 2005). Wooldridge in his book "Econometric analysis of cross section and panel data" (2002) provides this definition of endogeneous variable:

"An explanatory variable x_j is said to be endogenous in equation (3.1) if it is correlated with u."

In applied econometrics, endogeneity usually arises in one of four ways:

1. Omitted Variables

   Omitted variables bias appears in case, when it is required to control for one or more additional variables but, usually because of data unavailability, it is impossible to include them in a regression model.

2. Self-selection

   Self-selection bias can often cause correlation of explanatory variables with unobservables: if agents choose the value of x_j, this might depend on factors that are unobservable to the analyst.

3. Measurement error

   This is the case, when some of independent variables are measured with (a non-random) error. Assume the effect of a variable x_k^* is intended to be measured, but it is possible to observe only imperfect measure x_k. In the case, when the equation is estimated using x_k instead of x_k^* a measurement error is necessarily put into u.

4. Simultaneity

   Simultaneity arises when at least one of the explanatory variables is determined simultaneously along with y. If x_k is determined partly as a function of y, then x_k and u are generally correlated (J. Wooldridge, 2002).

3.2 Theoretical background of IV Regression

In case when endogenous variable is present in a model, what is needed is a method to generate only exogenous variation in x_K, which can obviously be done using a randomized experiment, but unfortunately for most economics applications such experiments are too expensive or simply not feasible. In this situation the instrumental variables estimator could provide a way to nonetheless obtain consistent parameter estimates. This method, widely used in econometrics and rarely used elsewhere, is conceptually difficult and easily misused.

The underlying population model is the same as in equation (3.1), but the unobservable error is explicitly allowed to be correlated with the explanatory variable. Let's assume, that variable x_K is endogenous. In order to use IV approach it is necessary to find explanatory variable denoted z_1, which has the property that changes in z_1 are associated with changes in x_K but do not lead to change in y (aside from the indirect route via x_K). This leads to the following path diagram:

$$\begin{tikzcd}
z \arrow[r] & x \arrow[r] & y \\
 & u \arrow[u] \arrow[ur] &
\end{tikzcd}$$

which introduces a variable z_1 that is causally associated with x_K but not u (Cameron and Trivedi, 2005).

In order to use the IV approach with x_K endogenous, an explanatory variable z_1 must satisfy the following **2 conditions**:

1) Instrument z_1 must be uncorrelated with the error term u :

$$COV(z_1, u) = 0 \tag{3.5}$$

2) Coefficient on z_1 in the linear projection of x_K onto all the exogeneous variables is nonzero:

$$x_K = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + ... + \delta_{K-1} x_{K-1} + \theta_1 z_1 + r_K \tag{3.6}$$

where, by definition of a linear projection error, $E(r_K) = 0$ and r_K is uncorelated with $x_1, x_2, ..., x_{K-1}$ and z_1

$$\theta_1 \neq 0 \tag{3.7}$$

When z_1 satisfies conditions (3.5) and (3.7), then it is said to be an instrumental variable (IV) candidate for x_K. The linear projection in equation (3.6) is called a reduced form equation for the endogenous explanatory variable x_K (J. Wooldridge, 2002).

The first assumption (3.5) excludes the instrument z_1 from being a regressor in the model for y. If it does not hold and y is correlated with both x_K and z_1, but regressed on x_K alone, than z_1 is being absorbed into the error term u and as a result is correlated with it. The second assumption (3.7) requires that there is some association between the instrument and the endogenous variable x_K, or in other words that z_1 is partially correlated with x_K once the other exogenous variables $x_1, ..., x_{K-1}$ have been netted out (Cameron and Trivedi, 2005).

The linear projection in equation (3.6) is called a reduced form equation for the endogenous explanatory variable x_K. In the context of single-equation linear models, a reduced form always involves writing an endogenous variable as a linear projection onto all exogenous variables. From the structural equation (3.1) and the reduced form for x_K, we obtain a reduced form for y by plugging equation (3.6) into equation (3.1) and rearranging:

$$y = \alpha_0 + \alpha_1 x_1 + ... + \alpha_{K-1} x_{K-1} + \lambda_1 z_1 + v \tag{3.8}$$

where $v = u + \beta_K r_k$ is reduced form error, $\alpha_j = \beta_J + \beta_K \delta_j$, and $\alpha_1 = \beta_K \theta_1$. According to assumptions (3.5) and (3.7), v is uncorrelated with all explanatory variables in equation (3.8), and as a result ordinary least squares (OLS) method consistently estimates the reduced form parameters, the α_j and λ_1 (J. Wooldridge, 2002).

Given a random sample $[(\mathbf{x}_i, y_i, \mathbf{z}_{i1}) : i = 1, 2, ..., N]$ from the population, the **instrumental variables estimator** of β is defined as:

$$\hat{\boldsymbol{\beta}} = \left(N^{-1} \sum_{i=1}^{N} \mathbf{z}_i' \mathbf{x}_i \right)^{-1} \left(N^{-1} \sum_{i=1}^{N} \mathbf{z}_i' y_i \right) = \left(\mathbf{Z}' \mathbf{X} \right)^{-1} \mathbf{Z}' Y \tag{3.9}$$

where **Z** and **X** are $N \times K$ data matrices and **Y** is the $N \times 1$ data vector on the y_i. The consistency of this estimator is immediate from the solution (3.10)

$$\boldsymbol{\beta} = [E(\boldsymbol{z}' \boldsymbol{x})]^{-1} E(\boldsymbol{z}' y) \tag{3.10}$$

and the law of large numbers (J. Wooldridge, 2002).

When searching for instruments for an endogenous explanatory variable, assumptions $COV(z_1, u) = 0$ (3.5) and $\theta_1 \neq 0$ (3.7) are equally important in identifying β. There is, however, one important difference between them: condition (3.7) can be tested, but condition (3.5) cannot, because the covariance in condition (3.5) contains the unobservable u and hence it is impossible to test fulfilment of this assumption (3.5). With regard to the second condition, it can be easily tested using the t test after OLS estimation, unfortunately r_K should not necessarily satisfy homoskedasticity assumption of OLS estimator, due to that fact a heteroskedasticity-robust t statistic for θ_1 is often recommended (J. Wooldridge, 2002).

3.3 Examples of IV regression in the literature

In this subsection some interesting examples of instrumental variables use from the literature will be provided.

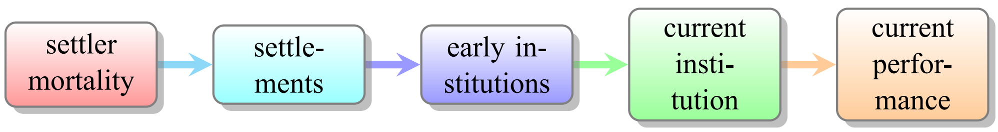
Settler mortality as an instrument for current institutions

Without doubt, one of such attractive cases is a study ”The Colonial Origins of Comparative Development: An Empirical Investigation” by Daron Acemoglu, Simon Johnson and James A (2001). Robinson. The authors are trying to find the answer to the fundamental question of economic science: what are the primary causes of large differences in income per capita across countries? The authors investigate differences in European settlers mortality rates to estimate the effect of institutions on economic performance.

In this paper, a theory of institutional differences among countries colonized by Europeans was proposed and exploited in order to derive a possible source of exogenous variation in institutions. In particular, Acemoglu, Johnson and Robinson pointed out, that the history of colonisation resulted in different institutions being formed: some countries received extractive institutions (whereby the coloniser would simply extract all resources but would not build proper institutions to promote growth and sustainable living) whilst others received inclusive institutions with strong emphasis on private property and checks against government power.

The colonization strategy was not random, the authors argue, that it was influenced by the disease environment not favorable to Europeans. Consequently, Europeans were unable to settle in variety of African territories,but Americas, Australia and New Zealand were successfully colonized.The authors suppose, that these institutions persisted even after independence. The author’s theory is schematically summarized on the following figure (3.1).

Figure 3.1: Connection between settler mortality, institutions and current performance in colonized countries



Source:Acemoglu et al., 2001

As was already mentioned earlier in the section (5.2) ,that in order to use IV approach it is necessary to find instrument z, which is correlated with endogeneous explanatory variable x_K, but is not correlated directly with dependent variable y (though z can be correlated with y indirectly through x_K). Fortunately, instrument settler mortality fulfills this property, because deceases, which caused high settler mortality had only limited impact on aboriginal adults, who developed various types of immunities. These diseases are therefore unlikely to be the reason why many countries in Africa and Asia are very poor today, hence the instrument indeed is not directly correlated with the dependent variable representing current performance, but only indirectly through institutions.

Minimal school-leaving age as instrument for cognitive skills

Another interesting case of instrumental variable use is paper ”Returns to skills around the world: Evidence from PIAAC”, written by Eric A.Hanushek,Guido Schwerdt,Simon Wiederhold and Ludger Woessmann in 2015. In this article the authors address the issue of the labor-market returns to human capital, they study, in particular, the impact of measured cognitive skills on the individual hourly wage. The data source used was the Programme for the International Assessment of Adult Competencies (PIAAC), developed by the OECD and collected in between 2011 and 2012. PIAAC was designed to measure key cognitive and workplace skills needed for individuals to advance in their jobs and participate in society, and provides internationally comparable data about adult populations in 24 countries.

The survey included an assessment of cognitive skills in three domains:

1. **Literacy**: Ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one’s goals, and to develop one’s knowledge and potential;
2. **Numeracy**: Ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life;
3. **Problem solving in technology-rich environments**: Ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks.

The authors consider 3 main types of possible bias in their analysis, in particular: measurement error, reverse causation and omitted variables.

1. **Measurement error** - implicit errors in measuring individual cognitive skills. The authors address possible attenuation bias by using literacy skills as an instrument for numeracy skills. They argue, that this approach essentially takes the variation that is common to both skill measures as the relevant cognitive dimension. The result of 2 Stage Least Squares (2SLS) estimation indicates, that literacy is a very strong instrument for numeracy.
2. **Reverse causation** - people may have better skills because they have a better job,in particular, better jobs may show tendency to sustain skill levels by regularly practicing skills or by providing the money to invest in adult’s education and training. In such case a variable, which is related to an individual’s skills but observed before the start of the labor-market career should be used. As a result the authors use Years of schooling or Parental education as instrumental variables.
3. **Omitted variables** - omitted variables related to skills and earning; for example, in the case, when family ties help people find a better job, the association between skills and earnings would be biased. In this instance, the authors argue, that family background should be used in the earnings equation as a control variable, rather than an instrument (Hanushek et al., 2015).

Other interesting cases

Instrumental variables regression can be also used in health economics. For example Evans and Ringel (1999), in their study examine the possibility of improving birth outcomes through higher cigarette taxes. In particular, the authors use cigarette tax as an instrument for maternal smoking.

Another interesting example of IV approach in literature is a study, written by McClellan et al. (1994), where authors examine the impact of more intensive hearth attack treatment on elderly mortality. Proximity to cardiac care centers was used as an instrument for heart attack surgery due to the fact, that treatment received by patients differed in observable and unobservable health characteristics, thus biasing estimates of treatment effects. Authors claim, that patients' distances to cardiac care centers are strong independent predictors of how intensively a patient with heart attack will be treated and appear uncorrelated with health status. Thus, differential distances approximately randomize patients to different likelihoods of receiving intensive treatments.

Part II

Practical part

Chapter 4

Data set for analysis of employee performance

In this chapter the data set used for analysis of employee performance is presented. The first section describes, how the data were acquired. In the second section main information about all variables as well as justification for their use is provided. And the third section of the chapter presents description statistics of data set together with graphical representation and further details about data and variables.

4.1 Data source

Internal individual level data about employees of real Prague insurance firm (later the Firm) in years 2018 - 2020 were used for the analysis. The majority of explanatory variables, for example variables representing employee's gender, experience in the Firm, education, nationality, age, employee's manager etc. as well as dependent variable measuring performance (explanations of all variables used in the analysis can be found down below in subsection 4.2) were taken from an internal Human resource (HR) database of the Firm. Working from home yearly statistics were taken from SAS database, where employees are required to note days, which were spent working from home or in other words "on Home office"(HO).

Distance from home to official workplace were calculated based on particular employee's home address and an address of his/her office, this data were acquired from HR internal database as well. Google maps platform, in particular The Distance Matrix API was used for precise distance calculation. This is a service which provides travel distance and time for a matrix of origins and destinations. The API returns information based on the recommended route between start and end points, as calculated by the Google Maps API, and consists of rows containing duration and distance values for each pair. For purpose of this master thesis, the distance values were chosen as more representative.

All back-office and middle-office employees with finalized Key Performance Indicators (KPI) evaluation were included in the data set. Data were merged together using Power Query tool in Microsoft Excel. As a result total number of 1363 observations were collected, in particular: 351 in year 2018, 477 in 2019 and 535 in 2020.

4.2 Description of variables

4.2.1 Dependent variable - Employee performance

Variable KPI_eval represents percentage fulfillment of individual Key Performance Indicators (KPI) at Year end evaluated by employee's direct manager, or in other words, variable illustrates performance of employee in particular year.

In the insurance company, where the data set were obtained from, individual KPI system works in this way:

1. At the beginning of year employee agrees with his/her manager on particular set of KPIs for the whole year, or in other words they agree on tasks, which employee should complete in one year. Part of KPIs could be for example successfull implementation of particular project(s), smooth quarterly/yearly reporting process, undergoing certain internal/external training's, developing a new skill, taking over new agenda etc. Normally there are around 6 or 7 KPIs and every one of them has weight (ranging from 0 to 100%), and the sum of all KPIs weights should be equal to 100%. Usually professional goals are the biggest part of KPI and their total weight oscillates around 40% -60%. Training and education KPIs for the most part weight from 20% to 40 % and the last 10% could be devoted to Soft KPI

2. During the year the manager is monitoring the performance of employee and fulfilment of his/her tasks.

3. After the end of particular year manager is required to evaluate employee's performance on every task and fill in a special KPI evaluation form. If employee successfully accomplished all tasks he/she receives 100% of bonus payment. If worker fulfills only part of KPIs or fulfills all, but the quality of execution is not good enough, he/she is going to receive corresponding percentage of bonus payment. It is apparent from the table 4.3, that sometimes final KPI evaluation could be higher than 100%. This is the case when employee outperforms his KPIs. In this instance, the worker could receive even more, than 100% of bonus payment, but it usually depends on the department's or HR's budget (Source: internal KPI methodology).

4.2.2 Explanatory variables

Home Office
It is numeric variable representing number of days which employee spent on Home office during a year.

Gender
Gender is dummy variable representing gender of the employees. It takes the value of 1 if the employee is woman and it takes the value of 0 if the employee is man.

Experience
Variable experience is numerical and represents the number of whole years since last date of entry in the Company till the time of performance evaluation (end of particular year).

Education
This variable represents the highest attained level of education at the time of performance evaluation (end of particular year). For better picture, the division into dummy groups is displayed in the following table 4.1.

Table 4.1: Variable Education

Level of education	Represented category	Number of observations
High school or lower	Primary, High school or Vocational education	476
Bachelor	Bachelor degree, Technical institute degree	152
Master	Master degree, Engineer degree	721
Doctor	Doctoral degree	11
NA's	Not listed in HR system	3

Source: the Firm's data processed in R software

Nationality
This is a categorical variable, which indicates employee's nationality. Macedonian, Croatian and Ukrainian categories were merged together with Slovak due to a very small number of observations (see table 4.2).

Table 4.2: Variable Nationality

Category	Nationality	Number of observations
Czech	Czech	1324
Slovak&Other	Slovak	34
	Macedonian	1
	Croatian	1
	Ukrainian	3

Source: the Firm's data processed in R software

Maternity leave
Maternity leave is a dummy variable, which takes the value of 1, if employee was on maternity leave during the specific year, nevertheless these people have also performance evaluation. There are only 7 such observations in the data set.

Year

This is categorical variable, which represents year and it takes values 2018,2019 and 2020.

Manager

This is a categorical variable, which represents a personal number of certain employee's direct manager, which makes yearly performance evaluation.

SmallChildren12 and SmallChildren6

These variables represent a number of certain employee's children, whose age was equal or smaller than 12 or 6 years accordingly at the time of performance evaluation (end of particular year).

Distance

This variable describes employee's distance from home to his workplace in kilometers.

dummy2020

It is dummy variable for year 2020, which is used for interaction with instrumental variables.

Age

This is numeric variable and it represents the age of employee at the time of performance evaluation (end of particular year).

4.3 Description statistics of dataset

In this section of the thesis description statistics of data set will be provided together with graphical representation and further details about data and variables.

4.3.1 Employee performance and Home office participation

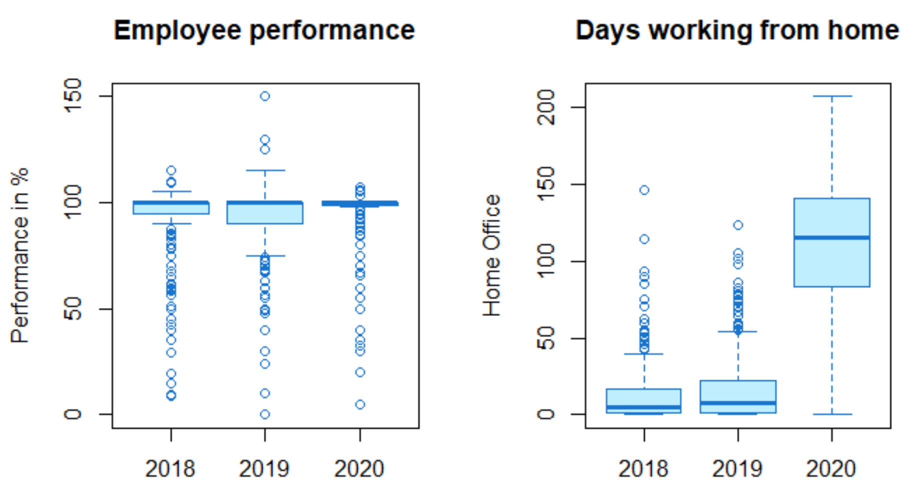
Table 4.3: Performance and Home Office statistics

	KPI_eval	HOYear
Min.	0.00	0.00
1st Qu.	95.00	4.00
Median	100.00	23.00
Mean	94.39	50.79
3rd Qu	100.00	101.00
Max.	150.00	207.00

Source: the Firm's data processed in R software

Figure 4.1 on the left shows the distribution of employee performance evaluation results in different years. The median of all 3 categories is equal to 100% . At first sight it seems, that employees' achievements are stable over 3 years and do not change significantly in 2020. As can be seen on the figure (4.1), there is almost no variability in employee performance evaluation in 2020 and median value is around 100%. It can be speculated, that during these difficult COVID-19 times, managers wanted to encourage theirs employees with good performance evaluation.

Figure 4.1: Boxplots of employee performance and Home Office participation



Source: the Firm's data processed in R software

Contrastingly, Home office participation increased drastically during the year 2020 as a result of COVID-19 preventive measures. Figure 4.1 on the right shows substantial increase in days worked from home during 2020. Median value of HO participation were only 5 and 8 days in 2018 and 2019 accordingly, but it reached 115 days during the year 2020.

Before the COVID-19 pandemic Home office was not very common in this insurance Firm and employees were using this possibility very rarely, even though the majority of workers had notebook for work and Home office could have been done without big technical issues. The situation changed in March 2020, working from home was strictly recommended during March, April, October and November, which was inline with governmental pandemic restrictions in Czech Republic (see Chapter 1). From May till September and also in December workers were allowed to come to the office, but management still encouraged employees to work from home in order to reduce spreading of the virus.

4.3.2 Distance

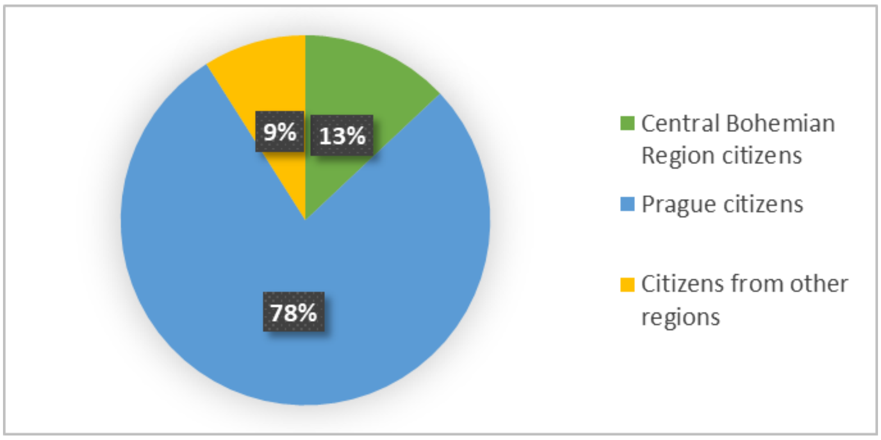
The information about the process of variable Distance calculation was already described in the section 4.1 Data source, and results can be seen on the figure 4.3 below. The first one represents all data set, big number of outliers is evident and values of Distance fluctuate from almost 0 to 500 km. Unfortunately the data obtained from The firm were not perfect, and some employees didn't have their actual address in HR database but the address of their permanent residence, for example an employee's parent's home in some town in Slovak republic. Such distance data cannot serve as successful instrumental variable, because it hardly could influence a decision making about Home office. Hence it was decided to remove outliers, but which value should be chosen as threshold?

Table 4.4: Distance statistics

	Distance	Distance (max 80km)
Min.	0.00	0.00
1st Qu.	7.20	6.30
Median	13.50	11.80
Mean	33.77	16.84
3rd Qu	30.77	31.70
Max.	507.00	77.60
NA's	5	0

Source: the Firm's data processed in R software

Figure 4.2: Job occupancy rate in Prague by regions



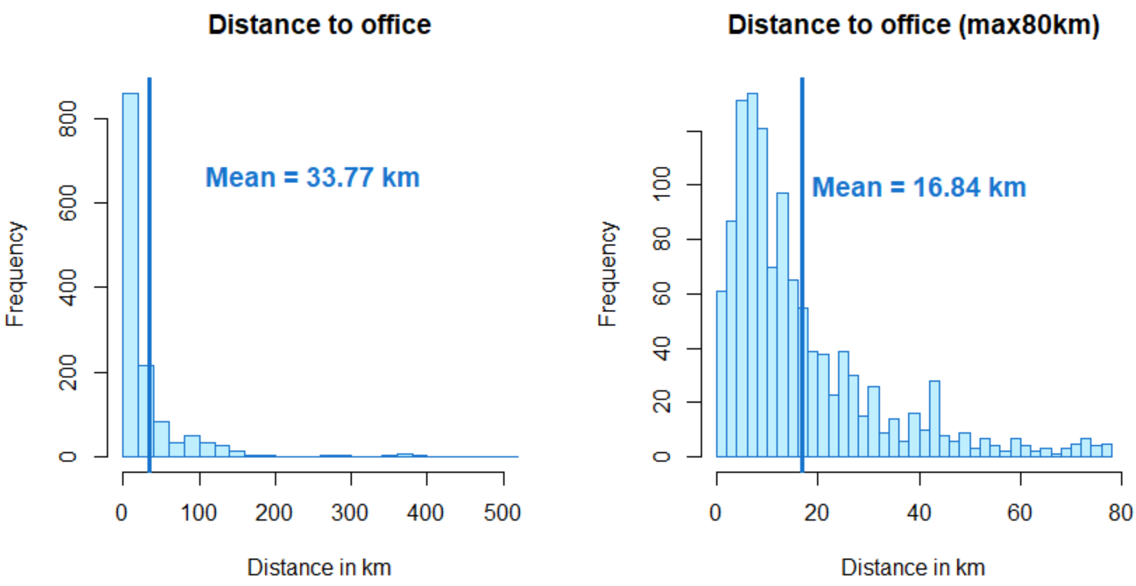
Source: "Dojížďka za prací a do škol v hl. m. Praze (na základě výsledků SLDB) - 2001", n.d.

As stated by Czech Statistical Office ("Dojížďka za prací a do škol v hl. m. Praze (na základě výsledků SLDB) - 2001", n.d.), 78% of people, who works in Prague are Prague citizens, the other 13% are commuting from Central Bohemian region and the rest are commuting from other regions. According to Google Maps, the distance between main towns of Central Bohemian region (such as Benesov, Beroun, Kolin Kutna Hora etc) and the Firm's office ranges from 60 to 80 km.

As a result, it was decided to chose threshold value 80km, which is supposed to cover job commuting from Central Bohemian region and Prague itself. Hence there are 1194 observation left in the data set (161 observations less). Different thresholds from 50 to 100 km were tried during the model estimation, but the results were in fact not significantly different. After all, only 100 observations out of 1358 have distance value ranging from 50 km to 100.

As can be seen on the table 4.4 below, there are some zero values of commuting distance as well. Probably these 7 people have their home address set as their working place, however, they still have reported their Home office in SAP database and the values are in normal range, hence it was decided to leave these observations in the data set.

Figure 4.3: Histograms of Distance



Source: the Firm's data processed in R software

4.3.3 Other variables

Table 4.5 on the right shows the descriptive statistics of explanatory variables Age and Experience. Median employee age in the insurance Firm is 41 years and mean age is almost 42. Half of people in data set are between 35 and 48 year old. With regard to variable Experience,which, as was already mentioned in the section 4.2.2 Explanatory variables, does not represent total professional experience of a person, but his/her experience in this particular firm, it's values range from 0 to 45(!) years, while median is equal to 7 years. 50 % of employees in data set have been working for this firm from 3 to 15 years.

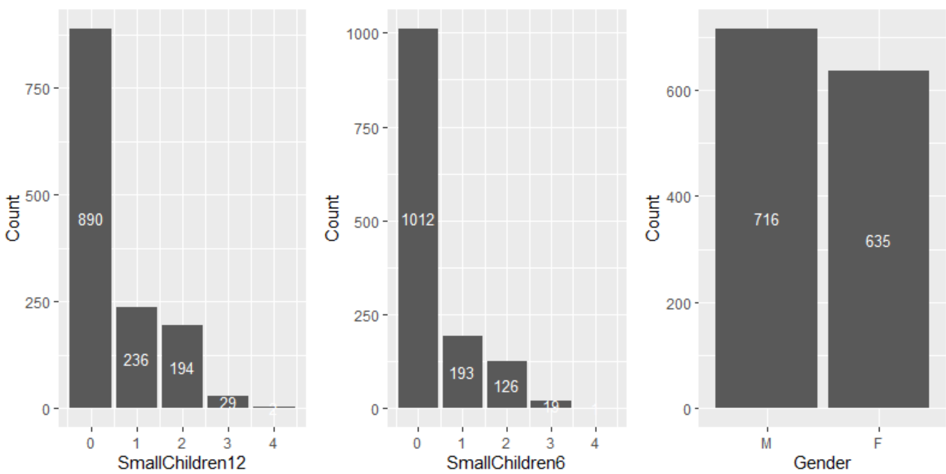
Table 4.5: Age and Experience statistics

	Age	Experience
Min.	21.0	0.0
1st Qu.	35.0	3.0
Median	41.0	7.0
Mean	41.9	9.5
3rd Qu	48.0	15.0
Max.	69.0	45.0

Source: the Firm's data processed in R software

On the graphs 4.4 variables Small Children and Gender are depicted on the X-axes against quantity of observations on Y-axis. Apparently there are 716 males and 635 females in the data set. With respect to variable Small Children, it can bee seen from the graph (4.4), that the majority of people in data set do not have any small children, and the rest have either 1 o 2, very small number of company's employees have more.

Figure 4.4: Bar charts of SmallChildren and Gender



Source: the Firm's data processed in R software

Chapter 5

Estimation of the model and results

In this chapter of the master thesis the results of the analysis are presented. The fists section describes the empirical strategy, the second one provides justification for instrumental variable use as well as presents the results of the First stage regression. The next two sections are dedicated to discussion of the results and also their possible limitations. Finally last section of this chapter explains, why Difference-in-Difference estimator was not used for analysis of employee performance.

5.1 Empirical strategy

The main explanatory variable Home office representing the number of days an employee spent working from home, is expected to be endogenous due to a possibility of self-selection bias: if agents choose to take Home office, this might depend on factors that are unobservable to the analyst. It is possible, that people, who expect to be more productive working from home, will chose to stay on Home office more often and otherwise. As a result self-selection bias can cause correlation of explanatory variables with unobservables and hence OLS will fail to provide consistent estimates of regression parameters.

Thus, it was decided to adjust for potential self-selection in Home office participation using a regression model with instrumental variables, in particular: Small Children, Distance and Age. Rationalization for use of these explanatory variables as instruments is provided in the next section (5.2).

In order to analyze the impact of working from home on employees' performance, 2 Stage Least squares estimator will be used due to the presence of possibly endogenous variable. This method is widely used in empirical economics research and is considered to be the most efficient instrumental variable estimator. It can be carried out "manually" using Ordinary Least Squares (OLS) with two-step procedure, in particular:

1. regress endogeneous variable x_K on instruments $1; z_1; ...; z_M$ together with control variables from the main regression and obtain the fitted values, say $\hat{x}_K$;

2. run the main OLS regression with fitted values $\hat{x}_K$ instead of original endogeneous variable x_K.

Nonetheless, in practice, it is best to use a software package with a 2SLS command rather than explicitly carry out the two-step procedure, due to the fact, that standard errors in second stage regression would be incorrect (J. Wooldridge, 2002). Finally, the main ”Second stage” regression model is presented below.

Second stage regression

$$\begin{aligned} Performance = \beta_0 + \beta_1 HomeOffice + \beta_2 Gender + \beta_3 Experience + \\ + \beta_4 Education + \beta_5 Nationality + \beta_6 MaternityLeave + \\ + \beta_7 Year + \beta_8 Manager + v, \end{aligned}$$

where v is an error term. The dependent variable in regression is Employee Performance, which represents percentage fulfillment of individual Key Performance Indicators (KPI) at year end evaluated by employee's direct manager, see section (4.2.1), where it is described in much bigger detail. The explanatory variable Home office indicates the total number of workdays in a year, which particular employee spent working remotely. The rest of the variables are controls for the socio-economic background of individuals.

According to the literature review, which was conducted in the chapter (2), it is expected that working from home or home office is on average positively associated with employee performance. Hence the hypotheses are:

H0: $\beta_1 = 0$

H1: $\beta_1 > 0$

It was decided to fit the model using 3 different estimators:

1. 2 Stage Least Squares (2SLS) under the assumption that variable Home office in endogeneous due to self-selection bias.
2. Ordinary Least Squares (OLS) assuming no endogeneity present in the model
3. Within or Fixed Effects estimator treating the given data set as unbalanced panel.

Fixed Effects estimator is used for analysis of panel data, which are *repeated observations on the same cross section, typically of individuals or firms in microeconomics applications, observed for several time periods* (Cameron & Trivedi, 2005). Fixed effects model allows for unobserved individual heterogeneity that may be correlated with regressors. Such unobserved heterogeneity leads to omitted variables bias that could in principle be corrected by instrumental variables methods using only a single cross section, but in practice it can be difficult to obtain a valid instrument (Cameron & Trivedi, 2005). The instrumental variables for Home Office (see 5.2), even though considered to be valid according to the Wald test (5.2), are not very strong either (robust F-statistics is lower than ”rule of thumb” value of 10). In this instance, Fixed effects estimator comes in handy.

The data set from the insurance Firm could actually be treated as short unbalanced panel, because the number of time periods T is not the same for all individuals. Within estimator offers an alternative way to proceed in these case under the assumption that unobserved individual-specific effects are additive and time-invariant. In a short panel the Within estimator measures the association between individual-specific deviations of regressors from their time-averaged values and individual-specific deviations of the dependent variable from its time-averaged value. This is done using the variation in the data over time (Cameron & Trivedi, 2005).

5.2 Instrumental Variables and First Stage regression

As it was already mentioned in the previous section, The main explanatory variable Home office might be endogenous due to a possibility of self-selection bias. Thus, it was decided to adjust for potential self-selection in Home office participation using a regression model with instrumental variables, in particular: Small Children, Distance and Age.

Firs stage regression

$$\begin{aligned} HomeOffice = {} & \beta_0 + \beta_1 Age + \beta_2 Age \times Dummy2020 + \beta_3 DummyChild + \\ & + \beta_4 DummyChild \times Dummy2020 + \beta_5 Distance + \\ & + \beta_6 Distance \times Dummy2020 + \beta_7 Gender + \beta_8 Education + \\ & + \beta_9 Nationality + \beta_{10} MaternityLeave + \\ & + \beta_{11} Year + \beta_{12} Manager + r, \end{aligned}$$

where r is an error term. Endogeneous variable Home Office indicates the total number of workdays in a year, which particular employee spent working remotely. Age, Small child dummy and Distance in kilometers are instrumental variables in the Fist stage regression. It was decided to include also interactions with dummy variable for the year 2020 due to the fact, that impact of instruments on decision making can be higher during the COVID-19 pandemic. The rest of variables are controls for the socio-economic background of individuals.

In regard to variable Distance, it is expected, that people, who live farther from office are more likely to work from home and otherwise, hence positive relationship with the variable Home Office is assumed.

As to instruments Small Children, also a positive relationship with Home Office is anticipated, since parent's decision weather to go to office may be influenced for example by sickness of his/her child etc. Moreover, during the year 2020 schools were closed due to COVID-19 pandemic, so parents of small children had almost no opportunity to go to the office. Due to that fact, the variable dummy2020 was also included, since an impact of instruments on decision making can be higher during the COVID-19 pandemic.

It was decided to use Small Children not as a numeric variable but as dummy, which takes the value of 1 once employee has at least one small child, which is more reasonable. It is expected, that a person becomes more likely to work from home once he/she has a small child, but significant difference in quantity of Home Office for a person with 1 small child and person with 2 is not too probable. The threshold for age of children was not taken arbitrary: age before 12 was chosen according to Ahrendt et al. (2020) and age before 6 represents preschoolers.

For sure there could be other unobservable variables correlated with employee's decision, weather to stay on Home office or no. For example, how much is person afraid of COVID-19: if employee is afraid of COVID, he/she is more likely to work from home. It would be good to use data about employee health status in this instance, unfortunately, such information is not collected by HR department, and thus, is unavailable for the analysis.

Table 5.1: First stage IV regression results (1)

	(1)	(2)	(3)	(4)
Home office				
(Intercept)	9.138	10.514(***)	13.393(**)	11.545(***)
Age	0.082		0.005	
Child 12 dummy	7.890(***)	7.828(***)	6.933(***)	7.172(***)
Distance (km)	0.201(**)	0.199(**)	12.884(*)	0.126(*)
Experience		0.185		0.189
GenderF	8.995(***)	8.776(***)	9.136(***)	8.870(***)
Educ. bachelor	-1.932	-1.397	-1.764	-1.366
Educ. master	-0.212	-0.085	-0.136	-0.031
Educ. doctor	9.014	9.174	8.926	9.077
Nationality Slovak	2.654	3.472	2.120	2.925
Mat. Leave	-21.381(*)	-21.466(***)	-20.937(*)	-20.973(*)
Year 2019	0.339	0.210	0.385	0.284
Year 2020	109.519(***)	100.38(***)	96.907(***)	96.919(***)
Manager dummies				
Child dummy :dummy2020	-2.1748	-1.467		
Distance :dummy2020	-0.1807(.)	-0.175		
Age:dummy2020	-0.2021			
F-test	32.44(***)	32.73(***)	32.88(***)	33.02(***)
R squared adj.	82.8%	82.9%	82.8%	82.8%

Source: the Firm's data processed in R software

Notes: The table demonstrates coefficients from the First Stage IV regressions estimated using Ordinary least squares OLS. Child 12 dummy means, that dummy variable for children under 12 was used in regression. Manager dummies are multiple dummies representing employees' direct managers, the coefficients were not included in the table due to large quantity of them. P values for t-test were calculated using the 2-way clustered robust standard errors, which were clustered by employee's personal number and his/her manager. *** denotes 0.1% significance, ** 1% significance, * 5% significance and (.) denotes 10% significance. F-test statistics as well as adjusted R squared are presented at the bottom of the table.

The second good predictor of employee risk aversion in regard to COVID-19 could be their age. Luckily, the variable Age is available in the data set and can be used as 3rd instrumental variable in the model. Age is highly correlated with Experience, hence it is reasonable to use only one of them in the regression, otherwise multicollinearity problem will emerge, which reduces the precision of the estimate coefficients and weakens the statistical power of regression model (J. M. Wooldridge, 2009).

There are 2 assumptions that must be satisfied in order to use explanatory variables as instruments (see section 5.2), particularly: explanatory variables should be correlated with the causal variable of interest (Home Office), but uncorrelated with any other determinants of the dependent variable (Performance) (Angrist & Pischke, 2009). The first condition can be tested easily using student t-test on corresponding β coefficient in a fist stage regression, but unfortunately second one cannot, and its fulfilment can only be judged hypothetically (J. Wooldridge, 2002).

Table 5.2: Linear Hypothesis Testing

IV model	Restictions	F stat.	Robust F stat.
Child 12 + Distance + Interactions	Interactions with dummy2020	1.87	1.23
Child 12 + Distance	Child 12 + Distance	12.66 (***)	6.44 (**)
Child 6 + Distance	Child 6 + Distance	9.52 (***)	4.335 (*)
Distance	Distance	6.48 (*)	2.33

Source: the Firm's data processed in R software

Notes: The table demonstrates results of comparing original IV model with restricted model using Wald test. Robust F statistics are calculated using Clustered Covariance Matrix of unrestricted model, which was clustered using employee's personal number and his/her manager. *** denotes 0.1% significance, ** 1% significance, * 5% significance and (.) denotes 10% significance.

The results of fist stage regression are presented in the table (5.1). Unfortunately, the instrument Age is not statistically significant, hence violating the first assumption for instrumental variables. Moreover, exclusion restriction may not be valid for this variable, as employee performance might increase with age. Consequently, it was decided to exclude Age from the list of instruments and use Experience as a control variable instead.

Furthermore, interaction members are neither significant, which can be interpreted as that impact of instruments on person's decision making didn't changed during the COVID-19 pandemic in comparison with previous years. This assumption can be further tested by means of Linear Hypothesis testing, specifically the original IV model will be compared with the restricted one and F-test will be used to assess whether they are significantly different from each other. The results of different F-tests are presented in the table (5.2), according to them the interaction members are indeed not statistically significant.

As a result of such analysis, the possibilities for suitable First Stage IV regression model are now narrowed down to 3 of them, which can bee seen in the table (5.3). The fist specification includes Distance and dummy for children younger than 12 years old as instruments, the second one contains the dummy for children under 6 instead and the last specification includes only Distance as instrument. All instruments are statistically significant, hence fulfilling the first instrumental variable assumption. Valid exclusion restriction is also assumed.

High adjusted R squared in these specifications is likely caused by very high number (more than 200) of dummy variables representing employee's managers. It is not really appropriate to estimate that much parameters using only 1190 total observations, but there isn't any other possibility. It is not reasonable to remove them either due to high statistical significance of these variables.

One possible solution would be to reduce quantity of manager dummies using a clustering method and group them according to their "strictness" or "liberality" and thus make clusters of managers, who give on average better performance evaluation to their employees, and clusters of managers, who give worse. Information about managers was not originally present in the data set, but it was possible to trace such data back, as managers also receive yearly performance evaluation from superiors. As a result, it was possible to cluster managers based on observable characteristics such as gender, age, experience, education, nationality and number of subordinates using the well-known K-means clustering algorithm. In the end such clusters wasn't significantly correlated to the dependent variable employee performance and thus wasn't included in the regression. Hence the final solution was to leave manager dummies in the model and use clustered robust standard errors, grouped by employee's Personal Number and manager.

Returning to First Stage IV regressions, coefficient estimates for all three models are not drastically different and are more or less in the same range. The values of adjusted R squared are as well similar and are approximately 83%. In the table (5.2) further Weak instrument tests for all 3 specifications are presented. According to Linear Hypothesis testing results, the instruments are indeed relevant in the models 1 and 2. F-statistics for both specifications are high and close to 10 or even higher. F-statistics of first model is 12.7, which exceeds "rule of thumb" value of 10 (Angrist & Pischke, 2009). Robust F-statistics is lower but still significant at 1% level, hence specifications 1 is chosen as final First stage IV regression model.

The results of the first regression suggest, that once employee has at least one child under 12, he/she on average spends 7 days more working from home than person without little children. Other interesting finding is, that females spend almost 9 days more on Home office during a year than men, which might be connected to the fact, that women are still expected to devote more time to housework and childcare than men (Feng and Savani, 2020; Gajendran and Harrison, 2007b). The empirical analysis also shows that Distance to work place is indeed positively associated with Home office. The corresponding coefficient is 0.126 and statistically significant at the 5% level. Unsurprisingly, the empirical analysis also indicates that employees of the Firm have spent on average almost 100 more days on Home Office than compared to year 2018.

In regard to coefficients for Maternity Leave and multiple Manager dummies, the parameters estimates are not going to be interpreted due to the fact, there are little number of observations for every of these groups (for example dummy Maternity Leave has only 7 observations), as a result, coefficient estimates might be deflected.

Table 5.3: First stage IV regression results (2)

	(1)	(2)	(3)
Home office			
(Intercept)	11.545(***)	12.256(***)	15.885(***)
Child 12 dummy	7.172(***)		
Child 6 dummy		6.547(*)	
Distance (km)	0.126(*)	0.127(*)	0.123(*)
Experience	0.189	0.188	0.156
GenderF	8.870(***)	8.729(***)	7.353(**)
Educ. bachelor	-1.366	-1.025	-0.519
Educ. master	-0.031	0.133	0.699
Educ. doctor	9.077	9.476	7.991
Nationality Slovak	2.925	2.144	1.120
Mat. Leave	-20.973(*)	-22.033(*)	-17.323(*)
Year 2019	0.284	0.102	0.098
Year 2020	96.919(***)	95.428(***)	96.558(***)
Manager dummies			
F-test	33.023(***)	32.788(***)	32.533(***)
R squared adj.	82.8%	82.7%	82.5%

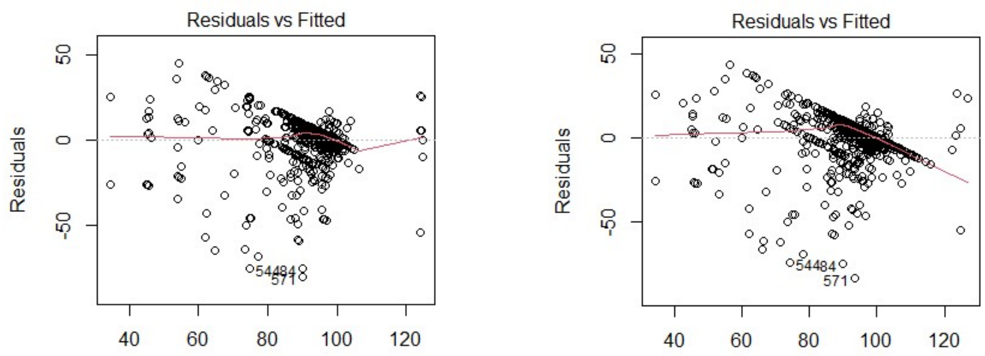
Source: the Firm's data processed in R software

Notes: The table demonstrates coefficients from the First Stage IV regressions estimated using Ordinary least squares OLS. Manager dummies are multiple dummies representing employees' direct managers, the coefficients were not included in the table due to large quantity of them. P values for t-test were calculated using the 2-way clustered robust standard errors, which were clustered by employee's personal number and his/her manager. *** denotes 0.1% significance, ** 1% significance, * 5% significance and (.) denotes 10% significance. F-test statistics as well as adjusted R squared are presented at the bottom of the table.

5.3 Results and discussion

This section provides the empirical results. The model was fitted using 3 different estimators: Stage Least Squares (2SLS), Ordinary Least Squares (OLS) and Within or Fixed Effects estimator (see 5.1). Table (5.5) shows the results of model estimation using three above mentioned methods: first two specifications are fitted with 2SLS estimator, specifications 3 and 4 wich OLS estimator, and finally the last two specifications fitted using Fixed effects. Presence of heterogeneity is dealt with using clustered robust standard errors grouped by employee Personal Number and his/her manager. Unfortunately, statistical software R does not enable clustering SE around other variables except for Year and Personal Number in panel regression model, hence SE for Fixed effects models are clustered by these variables.

Figure 5.1: OLS model and 2SLS model



The Table (5.5) shows, that specifications 1, 3 and 5 are statistically significant according to the Wald test, but there are actually no significant coefficients in the models 1 and 3 except for the intercept. In regard to Fixed effects specification (5), variables Experience, Education and Maternity leave are significant, but it is most likely caused not by real connection between these variables and employee performance but due to the very nature of Within estimator. For example, it seems like Experiance is positively associated with employee performance, but in reality it is probably the case, that people with high performance are staying in the company, as a result their experience is increasing. The similar situation is with education dummies.

Figure (A.1) depicts regression residuals against fitted values for OLS model on the left and for 2SLS model on the right. It can be seen from the figures, that residuals are not distributed randomly around the zero, and there is still some unexplained variance present in the model, which is probably due to missing unobservable variables. According to these 2 figures, OLS model looks more appropriate, as residual mean line is closer to horizontal zero line. According to Wu-Hausman formal endogeneity test (see table (5.4)), null hypothesis is not rejected, which means that the OLS and 2SLS estimators are not significantly different from each other. Such result means that OLS model is prefered in this case due to its efficiency (J. Wooldridge, 2002).

Table 5.4: Wu - Hausman test for Endogeneity

Model	statistic	p-value
2SLS	2.197	0.139
2SLS w/o Manager dummy	0.898	0.343

Source: the Firm's data processed in R software

As can be seen from the table (5.5), variable Home office is not significant in either of models, as a result null hypothesis is not rejected and positive impact of the home office on productivity has not been established. The results didn't fulfill the original expectations. Evidence from the literature suggests that in normal times people working from home can sustain, or even enhance, their productivity, while enjoying a better work life balance (Bloom et al., 2013; Rupietta and Beckmann, 2017). The possible reason for such outcome can be the fact, that the originally positive impact of Home office on employee performance was cancelled out due to the current pandemic situation. Stress caused by these difficult circumstances, lack of childcare, unsuitable working spaces etc. might be the reasonor not improved performance (JRC, 2020).

Another possible explanation to the question why the Home office coefficient isn't in line with theory, may be the fact, that there are almost no variability in performance evaluation in year 2020 (see figure (4.1)), which means that almost every employee received 100%. In order to test thus idea, the same regression model can be rerun using only data for the year 2018 and 2019, because there is at least some variability in performance evaluation during these years.

It was decided to estimate only OLS and 2SLS specifications, which are considered to be the best models. The table (5.6) shows, that the Home Office coefficient is still not significant in either of models. The possible explanation in this case could be the fact, that working from home during these years was extremely unpopular in the insurance Firm compared to the year 2020 (see figure (4.1)), and the median value on Home office participation was only 5 and 8 days in a year(!) accordingly, this is simply not enough working from home in order to have some influence on the performance.

Table 5.5: Results of model estimation

	2SLS	2SLS	OLS	OLS	FE	FE + IV
Employee performance	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	101.481(***)	93.195(***)	98.845(***)	92.491(***)		
Home Office	-0.143	-0.123	0.008	0.004	0.005	-0.407
GenderF	0.753	0.946	-0.331	0.569		
Experience	0.056	0.068	0.032	0.055	0.671(***)	1.447(*)
Educ. bachelor	-0.462	0.292	-0.365	-0.298	-7.242(*)	2.880
Educ. master	-0.245	1.509	-0.341	0.515	-6.553(*)	1.742
Educ. doctor	9.512(.)	4.283	8.3689(.)	2.354		
Nationality Slovak	2.434	-0.968	2.443	-1.003		
Mat. Leave	-6.498	-7.028	-3.859	-5.064	-10.250(*)	-17.747(.)
Year 2019	0.419	-0.447	0.388	-0.788	-0.298	-0.592
Year 2020	16.988	14.270	2.247	2.188	1.356	38.285
Manager dummies	yes	no	yes	no	no	no
	Chisq	Chisq	F-stat	F-stat	Chisq	Chisq
Wald-test	736.22 (***)	15.95	4.48(***)	1.64(.)	18.55(**)	11.40

Source: the Firm's data processed in R software

Notes: The table demonstrates coefficients from the final model estimated using OLS, 2SLS and Within estimator. Manager dummies are multiple dummies representing employees' direct managers, the coefficients are not included in the table due to large quantity of them. P values for t-test are calculated using the 2-way clustered robust standard errors, which were clustered by employee's personal number and his/her manager. For Fixed effects model, robust standard errors is clustered by employee's personal number and Year. *** denotes 0.1% significance, ** 1% significance, * 5% significance and (.) denotes 10% significance. Wald-test statistics are presented at the bottom of the table.

Table 5.6: Results of model estimation: years 2018 and 2019

	2SLS	OLS
Employee performance	(1)	(2)
(Intercept)	104.068(***)	100.588(***)
Home Office	-0.300	0.018
GenderF	-0.329	-1.756
Experience	0.042	-0.024
Educ. bachelor	-1.031	-1.371
Educ. master	-0.440	-0.467
Educ. doctor	12.342	12.083(*)
Nationality Slovak	1.152	2.176
Mat. Leave	-5.957	-3.895
Year2019	0.948	0.495
Manager dummies	yes	yes
Wald-test	3.288(***)	3.786(***)
R squared adj.	0,21	31.9%
R squared	34.3%	43.4%

Source: the Firm's data processed in R software

Notes: P values for t-test are calculated using the 2-way clustered robust standard errors, which were clustered by employee's personal number and his/her manager. *** denotes 0.1% significance, ** 1% significance, * 5% significance and (.) denotes 10% significance. Wald-test statistics are presented at the bottom of the table.

5.4 Possible limitations of the analysis

The fist limitation is that the results of the analysis have bounded external validity and may hold true only for a narrowly defined type of employee: middle office and back office of insurance and banking industry. In order to achieve higher level of external validity large data set with information about multiple firms from different industries is needed. The study from Rupietta and Beckmann (2017) is an example (see chapter (2)). However, one can argue, that the results of analysis in this master thesis should provide an useful insight even for many similar types of administrative positions, which are very often switched to the home-office environment during the COVID-19 pandemic.

Another issue could be the fact, that the data set provided by the insurance Firm is not perfect and includes noises. For example employees' addressees, which were used for calculation of distance to workplace, were sometimes not correct. In particular, some employees did not have their actual address in HR database but the address of their permanent residence, which could be totally different from the real place of residence. It was possible to remove some of incorrect distance values from the data set, by omitting the ones, which were obviously too high. But there was no way to disengage wrong addresses from correct once the distance value was lower than 80 kilometers threshold. If some employee moved from one part of Prague to another and did not report it to HR department, his distance to workplace value would be totally useless for estimating his HO participation.

The similar problem is with number of small children. Employees are providing the information about their children to HR department mainly in order to apply tax rebates, but only one of parents could use such discount. Hence it is probable, that the data set for analysis is actually missing a lot of records about employees' small children. As a result of such noises, the instrumental variables in 2SLS estimation are not strong, which could increase bias in the 2SLS estimator (it is never unbiased, but the weak instruments together with small data set could emphasizes large biases of 2SLS) (J. Wooldridge, 2002).

5.5 Analysis of employee performance using Dif-in-Dif

Difference-in-Difference (Dif-in-Dif) estimation has become an increasingly popular way to estimate causal relationships. Difference-in-Difference estimation consists of identifying a specific treatment or intervention (often the passage of law). The main idea of this method is comparison of the difference in outcomes before and after the treatment for affected samples with such difference for groups unaffected by it (Bertrand et al., 2003. The attractiveness of Dif-in-Dif estimation arises from its simplicity as well as its potential to overcome numerous endogeneity problems, which are common in case when comparisons are being made between heterogeneous individuals (Meyer, 1995, as cited in Bertrand et al., 2003).

Dif-in-Dif is typically estimated by Ordinary Least Squares (OLS) in panel or repeated cross-sectional data on individuals in treatment and control groups for several years before and after a specific treatment. Bertrand, Duflo and Mullainathan in their study ”How Much Should We Trust Differences-in-Differences Estimates” specified the relevant Dif-in-Dif regression model:

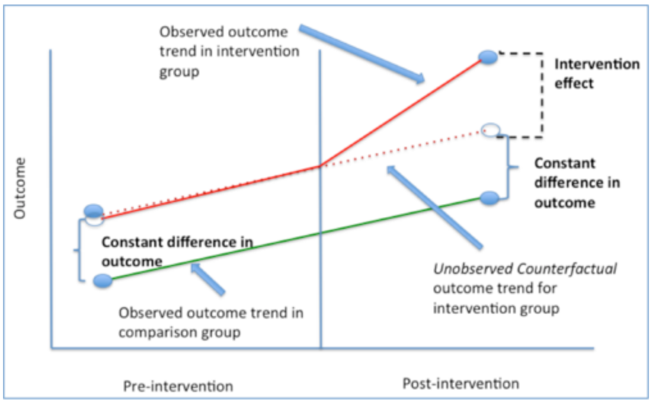
Let Y_{ist} be the outcome of interest for individual i in groups (such as a state) at time t and T_{st} be a dummy for whether the intervention has affected groups at time t:

$$Y_{ist} = A_s + B_t + cX_{ist} + \beta T_{st} + \epsilon_{ist} \tag{5.1}$$

where A_s and B_t are fixed effects for the states and years and X_{is} represents the relevant individual controls.

Thereby parameter $\hat{\beta}$ represents the estimated impact of the treatment. Standard errors around that estimate are OLS standard errors after accounting for the correlation of shocks within each state-year cell (Bertrand et al., 2003).

Figure 5.2: Difference-in-Difference estimation, graphical representation



Source:Health, n.d.

Two dimensions in Dif-in-Dif should not necessarily be states and time, although such example is archetypical in applied econometrics. Dif-in-Dif idea is much more general. Instead of states, the subscript st could indicate demographic groups, some of which are affected by the treatment and others are not. By the same token, instead of time, the data might be grouped by cohort or some other characteristic.

In order to estimate any causal effect using Dif-in-Dif approach, these four assumptions must hold:

1. Intervention unrelated to outcome at baseline (allocation of intervention was not determined by outcome)
2. the Parallel Trends Assumption, which states that the treatment group, absent the reform, would have followed the same time trend as the control group (for the outcome variable of interest)
3. Composition of intervention and comparison groups is stable for repeated cross-sectional design
4. The Stable Unit Treatment Value Assumption, which implies that there should be no spillover effects between the treatment and control groups, as the treatment effect then wouldn't be identified (Fredriksson and Oliveira, 2019)

In this diploma thesis Dif-in-Dif approach was supposed to be used for an identification of impact of working from home on employee performance during COVID-19 epidemic along with Instrumental Variable regression. It was intended to compare the difference in outcomes of treated and untreated groups, before and after the Stay-at-Home Policies were introduced in Czech Republic (year 2020). Employees with no possibility to perform their work remotely was going to serve as untreated or control group. And administrative workers, who were strictly advised by the government and the Firm to work as much as possible from home, were supposed to be the treatment group for Dif-in-Dif estimation. It was intended to use the percentage fulfillment of individual yearly Key Performance Indiocators (KPIs) as a proxy for employee performance of the Firm (same way as in Instrumental variable regression model).

Unfortunately it was not possible to acquire the data needed for control group. The only people, who cannot perform their work remotely are claims employees as well as agents. Claims employees file and take care of insurance claims, the part of their job is also to arrive at the place of the accident, check the car, if it is a car crush and so on. Such workers can do only part of their work remotely. Insurance agents are people, who deal with clients and sell insurance policies in particular. During 2020 this kind of employees could hardly perform work from home. Although the times has changed, and probably in the close future their work will be fully digitized and they will also be able to take Home office.

As was already said earlier, data about claims employees and agents unfortunately were not available in HR database, and even if they were, such group of workers would hardly serve as a good control group for analysis in this master thesis due to the fact, that claims people and agents are very different from administrative back-office and middle-office employees. Work of the first was in the end impacted by COVID-19 pandemic much more than the work of the later. For example, restrictions on movement of persons imposed by the government or the fact, that Firm's brunches were closed or had limited opening hours during the pandemic outbursts, affected agents performance significantly. Also pandemic restrictions could have had an impact on work of claims employees as well: less claims and accidents due to restrictions on movement of persons, and as a result less work. On the other hand, working day of back-office employee didn't change much at all, except for taking Home office more often than before.

One more problem is also the fact, that KPI systems work completely differently for claims employees and agents in comparison with the others: it depends in particular on the number of insurance policies sold or for claims - the number of times worker arrived to the place of accident and so on.

Due to all these reasons, the group of claims people and agents could not serve as appropriate control group for Dif-in-Dif estimation of Home office impact on the employee performance, even if such data were available. Moreover, it violates the third Dif-in-Dif assumption: *composition of intervention and comparison groups is stable for repeated cross-sectional design.* The negative impact of COVID-19 on work of agents and people working in claims, could have motivated them to leave the job, as a result the composition of control group would not be stable, as it is required by the assumption. Hence it was decided to use only Instrumental Variables regression for the analysis in this diploma thesis.

Conclusion

COVID-19 have shaped the world and remote working is going to stay with us at least in the form of hybrid working patterns. Companies are eager to improve cost-efficiency via reduction in office space, while providing employees with more flexibility in choosing where to work. In this instance, research on revealed employee performance while working from home is currently more scarce than ever.

The aim of this master thesis was to estimate the impact of working from home on employee performance during the COVID-19 epidemic. Based on the conducted literature research, it was expected, that working from home increases employee intrinsic motivation as well as his/her work satisfaction, and as a result leads to higher work performance (Hackman and Oldham, 1976; Rupietta and Beckmann, 2017; Bloom et al., 2013). But the result of model estimation did not support original hypothesis, and no impact of working from home on employee performance was found. The result was the same also in case of estimating the model only on data for years 2018 and 2019 (due to very low variability in performance evaluation in year 2020). The possible reason for such outcome can be the fact, that the originally positive impact of Home office on employee performance was cancelled out due to the current pandemic situation. Stress caused by these difficult circumstances, lack of childcare, unsuitable working spaces etc. might be the explanation why performance was not improved thanks to working from home (JRC, 2020).

Furthermore, some interesting finding in regard to home office participation were discovered during the analysis. For example, results of First stage regression show, that people with small children under 12 spend on average 7 more days on Home office during a year, than people without little children. The empirical analysis also shows that females spend almost 9 days more on Home office during a year than men, which might be connected to the fact, that women are still expected to devote more time to housework and childcare than men (Feng and Savani, 2020; Gajendran and Harrison, 2007b). The regression results also suggest that distance to work place is indeed positively associated with working from home.

My analysis is obviously not without weaknesses. One important drawback are noisy instrumental variables. In particular, employees' addressees, which were used for calculation of distance to workplace, were sometimes different from the real place of employees' residence. Part of the incorrect addressees was possible to disengage by omitting distance values higher than 80 kilometers. The data for second instrument representing small children were not perfect either. Employees are providing the information about their children to HR department mainly in order to apply tax rebates, but only one of parents could use such discount. Hence it is probable, that the data set for analysis is actually missing a lot of records about employees' small children.

Another weakness of this study is that the results of the analysis have bounded external validity and may hold true only for a particular type of back office employee from insurance or banking industry, and hence the findings cannot be generalized to a broader population. Therefore, further research is needed in order to analyze the impact of remote working on performance of employees from other economic industries or professions. Despite these weaknesses, I do believe that this study makes important contributions to the academic literature.

Bibliography

Torrance, E. P. (1962). *Guiding creative talent*. Prentice-Hall. https://doi.org/{10.1037/13134-000}

Hackman, J., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*(2), 250–279. https://doi.org/https://doi.org/10.1016/0030-5073(76)90016-7

McClellan, M., McNeil, B. J., & Newhouse, J. P. (1994). Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?: Analysis Using Instrumental Variables. *JAMA*, *272*(11), 859–866. https://doi.org/10.1001/jama.1994.03520110039026

Meyer, B. (1995). Natural and quasi-natural experiments in economics. *Journal of Business Economic Statistics*, *13*, 151–61. https://doi.org/10.1080/07350015.1995.10524589

Evans, W. N., & Ringel, J. S. (1999). Can higher cigarette taxes improve birth outcomes? *Journal of Public Economics*, *72*(1), 135–154. https://doi.org/https://doi.org/10.1016/S0047-2727(98)00090-5

Acemoglu, D., Johnson, S., & Robinson, J. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*. https://doi.org/10.1257/aer.91.5.1369

Harpaz, I. (2002). Advantages and disadvantages of telecommuting for the individual, organization and society. *Work Study*, *51*(2), 74–80. https://doi.org/10.1108/00438020210418791

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data (1st ed.)* The MIT Press.

Bertrand, M., Duflo, E., & Mullainathan, S. (2003). How much should we trust difference-in-difference estimators.

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2005). The Effects of Urban Spatial Structure on Travel Demand in the United States. *The Review of Economics and Statistics*, *87*(3), 466–478. https://ideas.repec.org/a/tpr/restat/v87y2005i3p466-478.html

Cameron, A., & Trivedi, P. (2005). *Microeconometrics: Methods and applications*. https://doi.org/10.1017/CBO9780511811241

Gajendran, R., & Harrison, D. (2007a). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *The Journal of applied psychology*, *92*, 1524–41. https://doi.org/10.1037/0021-9010.92.6.1524

Gajendran, R., & Harrison, D. (2007b). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences.

The Journal of applied psychology, *92*, 1524–41. https://doi.org/10.1037/0021-9010.92.6.1524

Mello, J. A. (2007). Managing telework programs effectively. *Employee Responsibilities and Rights Journal*, *19*, 247–261.

Angrist, J., & Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion.

Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach*. South-Western. http://books.google.ch/books?id=64vt5TDBNLwC

Rosenman, R., Tennekoon, V., & Hill, L. (2011). Measuring bias in self-reported data. *Int. J. of Behavioural and Healthcare Research*, *2*, 320–332. https://doi.org/10.1504/IJBHR.2011.043414

Dutcher, G. (2012). The effects of telecommuting on productivity: An experimental examination. the role of dull and creative tasks. *Journal of Economic Behavior Organization*, *84*, 355–363. https://doi.org/10.1016/j.jebo.2012.04.009

Bloom, N., Liang, J., Roberts, J., & Ying, Z. (2013). Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics*, *130*. https://doi.org/10.1093/qje/qju032

Demetriou, C., Özer, B., & Essau, C. (2015). Self□report questionnaires. https://doi.org/10.1002/9781118625392.wbecp507

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from piaac. *European Economic Review*, *73*, 103–130. https://doi.org/https://doi.org/10.1016/j.euroecorev.2014.10.006

Rupietta, K., & Beckmann, M. (2017). Working from home: What is the effect on employees' effort? *Schmalenbach Business Review*, *70*. https://doi.org/10.1007/s41464-017-0043-x

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., & Van Reenen, J. (2019). What drives differences in management practices? *American Economic Review*, *109*(5), 1648–83. https://doi.org/10.1257/aer.20170491

Fredriksson, A., & Oliveira, G. (2019). Impact evaluation using difference-in-differences. *RAUSP Management Journal*, *54*, 519–532. https://doi.org/10.1108/RAUSP-05-2019-0112

Monteiro, N. P., Straume, O. R., & Valente, M. (2019). *Does Remote Work Improve or Impair Firm Labour Productivity? Longitudinal Evidence from Portugal* (CESifo Working Paper Series No. 7991). CESifo. https://ideas.repec.org/p/ces/ceswps/_7991.html

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. https://doi.org/https://doi.org/10.1016/j.jpubeco.2020.104245

Ahrendt, D., Mascherini, M., Leoncikas, T., & Sandor, E. (2020). Living, working and covid-19: First findings – april 2020.

Bao, L., Li, T., Xia, X., Zhu, K., Li, H., & Yang, X. (2020). *How does working from home affect developer productivity? – a case study of baidu during covid-19 pandemic*.

Chang, Y., Chien, C., & Shen, L.-F. (2020). Telecommuting during the coronavirus pandemic: Future time orientation as a mediator between proactive coping and perceived work productivity in two cultural samples. *Personality and Individual Differences*, *171*, 110508. https://doi.org/10.1016/j.paid.2020.110508

Could czech's measure to fight coronavirus save thousands of lives? (2020). https://praguemorning.cz/could-czechs-measure-to-fight-coronavirus-save-thousands-of-lives-2/

Etheridge, B., Wang, Y., & Tang, L. (2020). *Worker productivity during lockdown and working from home: evidence from self-reports* (ISER Working Paper Series No. 2020-12). Institute for Social and Economic Research. https://ideas.repec.org/p/ese/iserwp/2020-12.html

Eurostat. (2020). Employed persons working from home as a percentage of the total employment, by sex, age and professional status. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_ehomp&lang=en

Fana, M., Milasi, S., Napierala, J., Fernández-Macías, E., & Vázquez, I. G. (2020). *Telework, work organisation and job quality during the covid-19 crisis: A qualitative study* (JRC Working Papers Series on Labour, Education and Technology No. 2020/11). Seville, European Commission, Joint Research Centre (JRC). http://hdl.handle.net/10419/231343

Feng, Z., & Savani, K. (2020). Covid-19 created a gender gap in perceived work productivity and job satisfaction: Implications for dual-career parents working from home. *Gender in Management: An International Journal*, *ahead-of-print*. https://doi.org/10.1108/GM-07-2020-0202

Green, N., Tappin, D., & Bentley, T. (2020). Working from home before, during and after the covid-19 pandemic: Implications for workers and organisations. *New Zealand Journal of Employment Relations*, *45*. https://doi.org/10.24135/nzjer.v45i2.19

JRC. (2020). Telework in the eu before and after the covid-19. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc120945_policy_brief_-_covid_and_telework_final.pdf

Purwanto, A., Asbari, M., Suryani, P., Cahyono, Y., Fahlevi, M., & Mufid, A. (2020). Impact of work from home (wfh) on indonesian teachers performance during the covid-19 pandemic : An exploratory study. *29*, 6235–6244.

Sostero, M. (2020). Teleworkability and the covid-19 crisis:a new digital divide? *A Joint European Commission–Eurofound Repor*, *JRC120945*.

Sostero, M., Milasi, S., Hurley, J., Fernández-Macías, E., & Bisello, M. (2020). *Teleworkability and the covid-19 crisis: A new digital divide?* (JRC Working Papers Series on Labour, Education and Technology No. 2020/05). Seville, European Commission, Joint Research Centre (JRC). http://hdl.handle.net/10419/231337

Ipsen, C., Veldhoven, M., Kirchner, K., & Hansen, J. (2021). Six key advantages and disadvantages of working from home in europe during covid-19. *International Journal of Environmental Research and Public Health*, *18*, 1826. https://doi.org/10.3390/ijerph18041826

Measures adopted by the czech government against the coronavirus. (2021). https://www.vlada.cz/en/media-centrum/aktualne/measures-adopted-by-the-czech-government-against-coronavirus-180545/

Ministerstvo vnitra české republiky. (2021). https://www.mvcr.cz/clanek/vlada-na-tri-tydny-zakazala-pohyb-mezi-okresy-omezeni-ma-zabranit-sireni-agresivnich-mutaci-koronaviru.aspx

Update 1-czech republic detects uk coronavirus variant, to maintain lockdown measures. (2021). https://www.reuters.com/article/health-coronavirus-czech-idUSL1N2JT1GH

Vyas, L., & Butakhieo, N. (2021). The impact of working from home during covid-19 on work and life domains: An exploratory study on hong kong. *Policy Design and Practice*, *4*(1), 59–76. https://doi.org/10.1080/25741292.2020.1863560

Arcgis. (n.d.).
Covid-19: Onemocnění aktuálně mzčr. (n.d.).
Dojížďka za prací a do škol v hl. m. praze (na základě výsledků sldb) - 2001. (n.d.). https://www.czso.cz/csu/czso/13-1127-04-sldb_2001-4__dojizdka_za_praci
Health, C. P. (n.d.). *Difference-in-difference estimation*. https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation
Advice for the public on covid-19. (March 2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public
Rogers, K. (March 2020). Is there a cure for coronavirus? https://www.britannica.com/story/is-there-a-cure-for-coronavirus
V české republice jsou první tři potvrzené případy nákazy koronavirem. (March 2020). https://koronavirus.mzcr.cz/v-ceske-republice-jsou-prvni-tri-potvrzene-pripady-nakazy-koronavirem/
Who director-general's opening remarks at the media briefing on covid-19 - 11 march 2020. (March 2020). https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

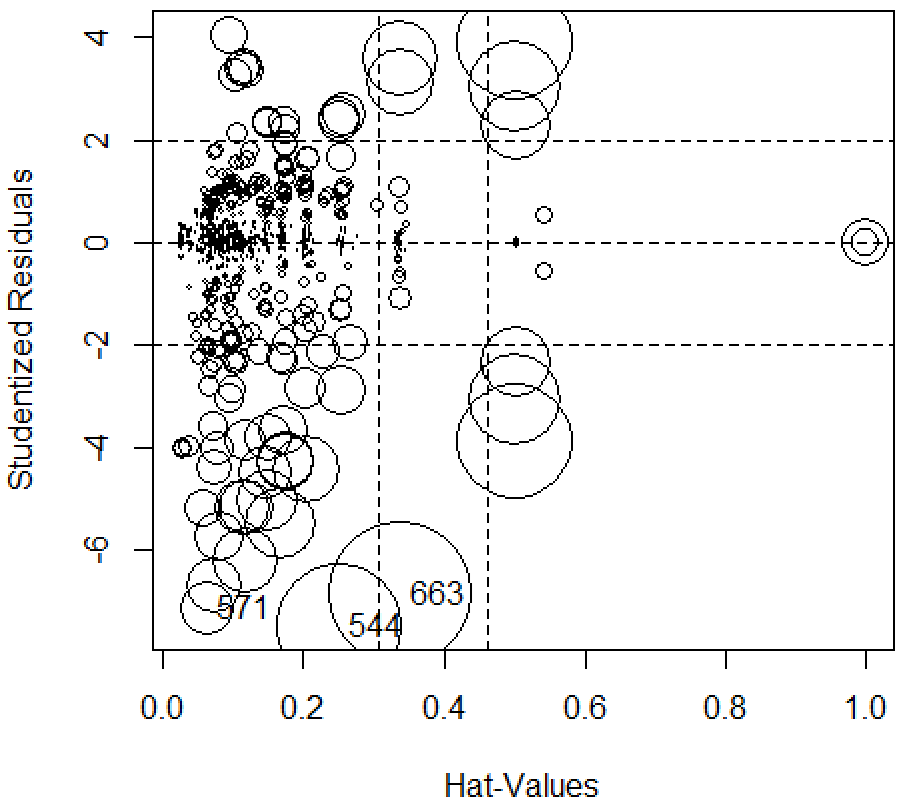
List of Tables

List of Figures

Appendix A

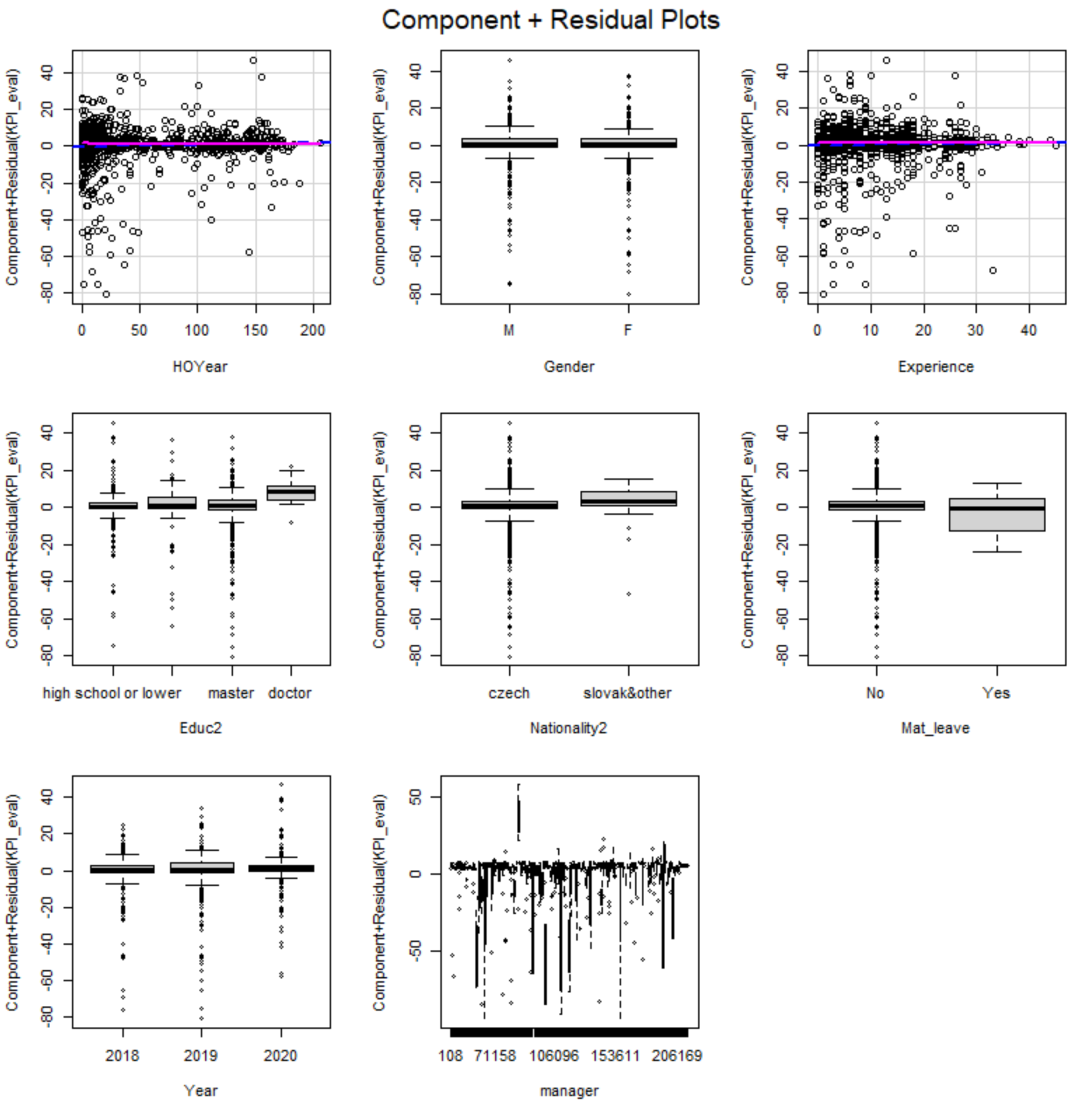
Appendix

Figure A.1: Hat values for OLS model (3)



Source: the Firm's data processed in R software

Figure A.2: Component + Residual (Partial Residual) Plots for OLS model (3)



Source: the Firm's data processed in R software