

Work Stress*

Sagith Kalaichelvam

Youngho Kim

16 March 2023

Abstract

This paper discusses data from some survey, which can gather data that is linked to how people face work stress. It uses the original data from the survey which determines the relationships between respondent work stress levels by satisfaction, status, sectors, and work time. There will be some strong connections between the analysis and figures of the work stress. This paper will provide a strong analysis about the work stress that people are facing.

Table of contents

1	Introduction	2
2	Data	2
2.1	Data Source and Collection	2
2.2	Data Cleaning and Modification	3
2.3	Data Visualization and Analysis	4
3	Results	8
4	Discussion	9
4.1	Work Stress Levels	9
4.2	Work Hours of a Work Shift & Weekly Schedule	9
4.3	A Potential Explanation: Shifting Priorities	9
4.4	Weakness and Moving Forward	10
4.4.1	Bias and Ethical Concerns	10
4.4.2	Data Collection Concerns	10
4.4.3	Moving Forward	10
	References	11

*Code, data are available at: <https://github.com/YoungKim164/workstress> and <https://doi.org/10.1257/jep.36.1.151> and replication package at: <https://doi.org/10.48152/ssrp-cdf6-2r92>

List of Figures

1	Work Stress by Job Status	4
2	Work Stress by Work Hours	5
3	Work Stress by Industry Sector	6
4	Work Stress by Job Satisfaction	7

List of Tables

1 Introduction

Work is the most common cause of many people to get stress. Work stress can be found in many forms for people. In this paper, we will examine job status, work hours, industry sector, satisfaction rate of your job into the correlation of work stress levels. There will be context that will motivate such as the data that is representing the factors of work stress and how people feel based on the real dataset. There are many analysis to talk about relevant to work stress, that can later conclude the next steps that workplace should consider to reduce work stress for many employees. All of the data are being collected through GSS, which allowed us to create different figures of job status, work hours, industry sector, and satisfaction rates. They were all written as code within the Quarto document. We saw some data that shows how the work stress occurs from the these factors. All of these information within the paper is very important because it will let the readers and others in the public to be aware of how people are facing work stress, and the analysis can provide some strong points that can lead everyone to know what employers and managers can do to avoid work stress for all employees working.

In this paper, we gathered data from (Publisher Name), (Data Name). With the gathered data, we generated figures to observe the trend in work stress of different categories using R (R Core Team 2020). We also uploaded and cleaned data using R (R Core Team 2020), tidyr (Wickham 2021), dplyr (Wickham et al. 2021), tidyverse (Wickham et al. 2019), and haven (Wickham and Miller 2021) packages. Figures and tables were created with knitr (Xie 2014), ggplot2 (Wickham 2016), kableExtra (Zhu 2021), and dplyr (Wickham et al. 2021).

In section 2 we discuss the source of the data used in this paper, the strengths and weaknesses the source and methodology contains, and potential blind spots that the data misses. Section 3 discusses the results that is being produced from the figure and what it tells us by its visualization. Section 4, we would provide a brief discussion about what we get from all figures, and make some thoughts. In section 5 there is an optional survey used to gather more information about stress caused by job status, work hours, industry sector, satisfaction rate of your job into the correlation of work stress levels.

2 Data

2.1 Data Source and Collection

The dataset for this paper were obtained from the US General Social Survey [] from NORC at the University of Chicago. The dataset was published in the year 2021. The raw dataset includes data from 568 Survey Questions for 4032 Survey Participants. GSS monitors pubic opinion and behavior in the United States of America. This survey has been conducted since 1972 by the NORC at the University of Chicago and funded by the National Science Foundation(NSF). We have selected 10 items from the raw dataset, and only extracted selected variables. These selected variables were imported in R using the package haven(Wickham and Miller 2021) and readxl(citereadxl?). Using the R(R Core Team 2020) package tidyverse(Wickham

et al. 2019) and dplyr(Wickham et al. 2021), we cleaned and performed exploratory data analysis on the selected dataset to get insights into the data. The insights gained from this data set is shown throughout all of the figures we created in this paper. The some of the selected items are year, travel time to work, labor force status, number of hours worked, and etc. The analysis of these collected information is available on section 2.3 Data Visualization and Analysis.

2.2 Data Cleaning and Modification

The data used in this paper were cleaned to get an accurate representation and remove any unnecessary information in this paper. Data were cleaned to remove data of some of results that does not have responds on specific items that we are looking into. Some of participants refused to answer or skipped some of the questions for various reasons. We filtered out data that would not have meaningful result on the graphs.

For each figures, we selected certain items that are closely related to work condition and explored how they affects the work stress. The Data set is divided into categories and range, like working 20 or less, 20-29, 30-39, 40-49, 50-59, 60 over hours of working per week, or parttime/fulltime worker. Eachfigures explores different categorization to see how they effect personal stress of work.

2.3 Data Visualization and Analysis

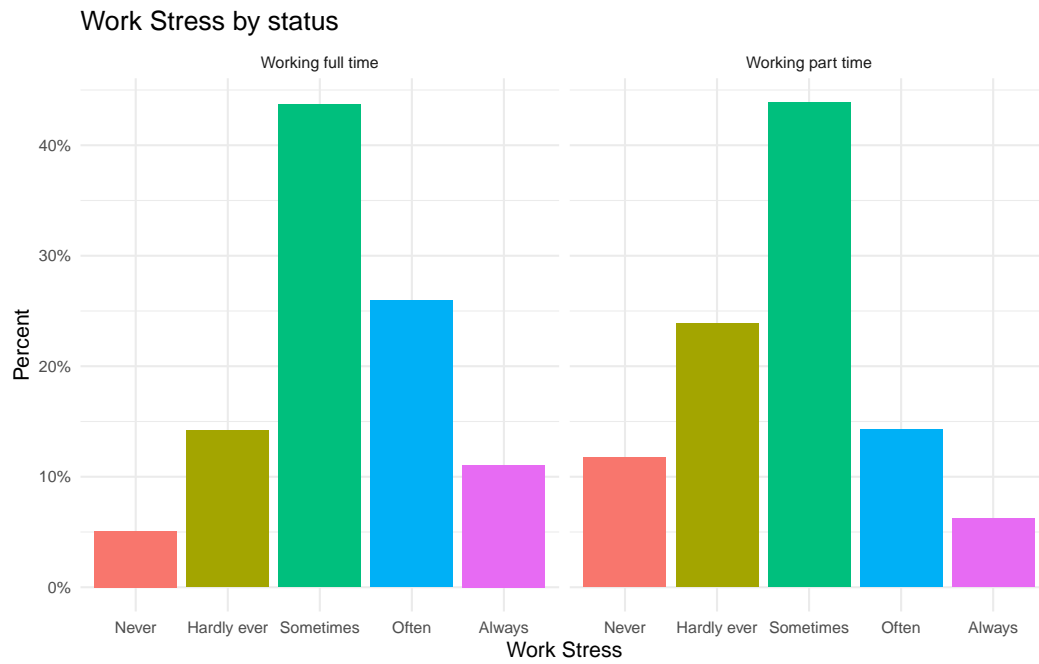


Figure 1: Work Stress by Job Status

Figure 1 shows the work stress by work status.

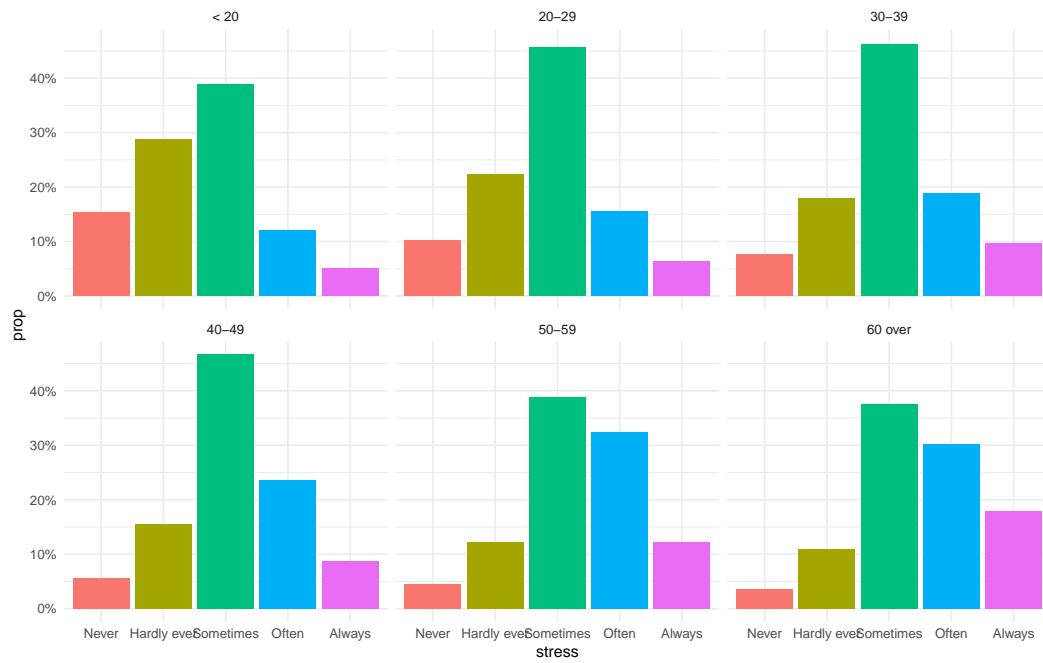


Figure 2: Work Stress by Work Hours

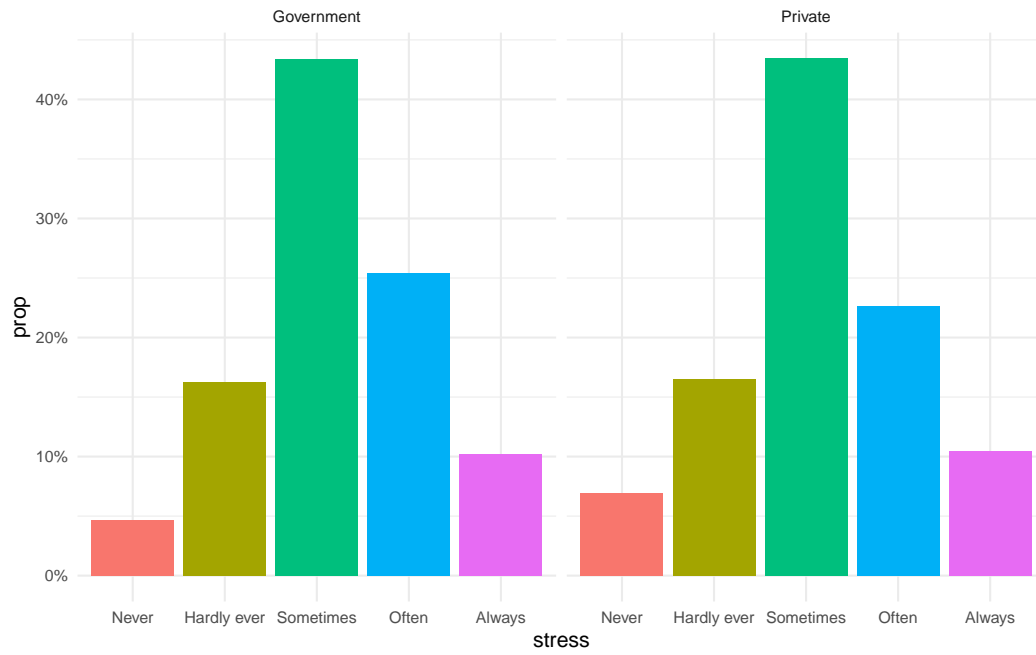


Figure 3: Work Stress by Industry Sector

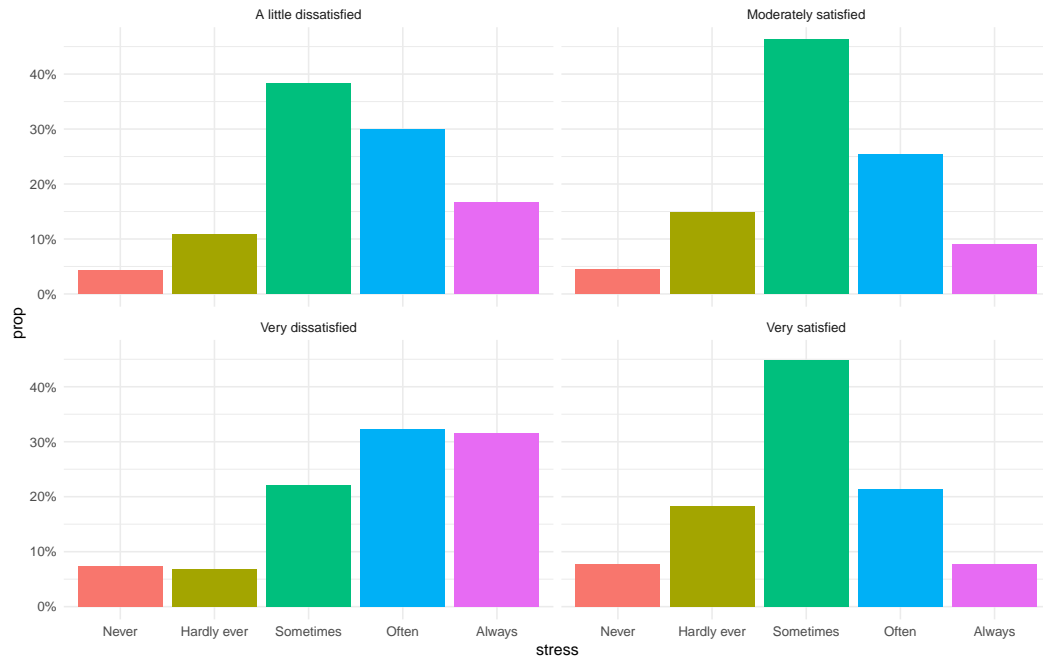


Figure 4: Work Stress by Job Satisfaction

3 Results

Figure 1 shows that over time, the number of births in the United States is dropping steadily, while it slightly went up on peak around 1990, 2001, and 2007. Most likely, the hikes going up during these years might have been into other factors that made more people among the selected population have kids. After 2007, the number of births drastically dropped a lot. Nowadays, among 1,000 women from age 15 to 44, only about 56 births will occur. While in the past, around the 1980s and 1990s, the number of births was around the 65 to 70 range. The line in the graph shows that in the near future, there will be a much lower amount of people having kids, and this can be due to other factors from their lifestyle, which is making them decide not to have more kids.

While the figure shows the birth rate trend is going down, there are certain spikes in the peak of birth rates around 1990, 2001, and 2007. There were some incidents that must have happened in the United States in those years, which may have caused people to stay home more often. As we know that people that are likely to be unemployed tend to have more kids as they will have more freedom during their time spent at home. This allows the birth rates to be higher than usual. During these years, Americans faced lower employment opportunities within society (Agency 2008). They do not have access to useful resources that improve their lifestyle in these years, so they are at home most of the time, which encourages couples to have an interest in having more kids, which ends up increasing the birth rates for these years.

4 Discussion

4.1 Work Stress Levels

Even though work stress have different causes. We believe that the major cause or initial step into work stress would traffic delays. There are always traffic delays that most commuters would face almost every day. Most of these commuters are people going for work especially during rush hour time. Satisfaction rate is very important for jobs, and most people are sometimes not satisfy, its mainly due to these traffic delays. For example, if you are taking Greater Toronto Area commuters, most jobs are found in downtown Toronto, North York, and Mississauga. The GTA is going far west as Hamilton to east as Durham Region. Most of these commuters will be spending at least 1 hour of commuting to their workplace either by car or public transit. In an anonymized therapy session in 2014, a client expressed: “by the the time I get home from work, I’m just, you know, after that drive, I’m just so dead, it’s like, you know, shaking and don’t want to drive anymore. I don’t want to do anything else. . . sometimes I like if there’s been lots of traffic or like if I’ve been tired and like been forcing myself to stay awake so I can get home.” Traffic is the main issue that we believe that is impacting the job satisfaction rate, which is why lots of people are facing work stress. We need to find a solution that can reduce the traffic, increase job satisfaction, and if possible these can expect to reduce the work stress for everyone.

4.2 Work Hours of a Work Shift & Weekly Schedule

Right now most professionalists are working 40 hours a week, 8 hours shift each day. There is a plan of increasing work hours to 10 hours shift, and work only 4 days each week. But, based on some data, it is definitely not recommended to increase work hours because currently lots of people are always facing work stress when the number of hours are higher. So, if the work shift gets extended longer to reduce the total number of work days, this may cause more work stress because lots of people will be stress over number of tasks to complete during work. Other alternatives that can be used to reduce work stress can be reducing the total number of hours per week from 40 hours to 35 hours. This will allow people to face less work stress because it will reduce the number of hours for them working, and we believe 7 hours work shift can be enough to complete important tasks in a job.

4.3 A Potential Explanation: Shifting Priorities

When it comes to the age of the population of the United States and comparing it from the past to now. It shows that more people tend to have kids much later than before. While fewer people have kids earlier than in the past. In the past, more people tend to have kids much earlier. There are several reasons why more people had kids at an early age back then; people used to get married much earlier, and many people were unemployed or doing a job for self-employment, which was not even an educated or professional job. The modern lifestyle that has been adopted in the past few decades tends to make people focus on all other areas of life, which makes them marry much later, and have kids later. Nowadays, many marriages happen at the age of late 20s or early 30s, while they tend to have their first child around the early to mid-30s. **?@fig-5** depicts the birth rates in women between the ages 30-39 and 35-39 positive slope trend. Whether pre or post-2007 recession, this trend looks steady over the past 40 years. This finding aligns with the current common belief and the social norm where, perhaps, women’s peak childbearing ages have shifted from the early twenties to early thirties. **?@fig-6** shows the average number of children born to women by age in five-year intervals from the 1968-1972 through 1993-1997 birth cohorts. These birth cohorts include women entering their peak childbearing years (20–24) in 1992, 1997, 2002, 2007, and 2017. This chart indicates that the three cohorts of young adult women born between 1968 and 1982 in 1992, 1997, and 2002 had similar childbearing age profiles. The 2007 birth cohort of women had fewer children in their 20s and 30s. The 2012 and 2017 birth cohorts (born between 1988 and 1997) have significantly fewer children than the preceding cohorts. This finding is consistent with a decline in births, primarily to more recent cohorts of women. From

a statistical point of view, this divergence across cohorts explains the sizable decrease in annual birth rates that began in 2007.

4.4 Weakness and Moving Forward

4.4.1 Bias and Ethical Concerns

As we know that this data was collected and created by other human beings, just as how we are. There are some ethical concerns that could affect the accuracy of our data because humans can make mistakes, and they can make decisions based on their preferences, which can affect the whole dataset. So, this analysis is only valid based on our dataset, which can be modified if the original dataset is changed.

The original data shows different races, such as White, Black, and Hispanic. They have included Hispanic based on all colours, which erases the identity of Black and Indigenous Hispanic people. This tends to be a bit of racism because the information provided on the data did not specifically indicate whether or not the Hispanic category includes everyone that identifies as Hispanic or is it just people that look Hispanic.

4.4.2 Data Collection Concerns

This dataset tends to collect huge data information on the population among people from the United States and considers lots of factors from race, birth rates by age, student debts, etc. Having too much information means that when the data is collected, there is a possibility of human error, calculation error, and chances of small modification of data due format of reporting an error. So, the way how the data was collected can have lots of problems because people can make mistakes easily.

4.4.3 Moving Forward

This paper raised concerns about a dramatic decline in U.S. birth rates starting in 2007. Through the paper, we proposed a general explanation for the decline in births across recent cohorts of U.S. women that focuses on the shifting priorities of cohorts. As dictated by the recent socio-phenomena that encompasses preferences for having children, life aspirations, and the nature of parenting, a change in the relationship between education and childbearing, a rise in childlessness, and the establishment of a two-child norm for those having children have irreversibly changed (Kearney, Levine, and Pardue 2022). This new social phenomenon that might be causing the dramatic decline in U.S. birth rates is important from an economic policy standpoint to acknowledge that an aging population and shrinking workforce pose challenges to economic growth and the sustainability of social insurance systems (Yang and Morgan 2003). Thus, the most appropriate way to address declining U.S. birth rates may be required: 1. developing A.I. and robots where a shrinking workforce poses a problem, 2. an appropriate policy to address issues with Social Security and Medicare on a secure basis for the long-term. Suppose the U.S. birth rate is not showing any reverse trend shortly. In that case, the U.S. economy and political system should enforce greater emphasis on addressing the issue before it becomes irreversible.

References

- Agency, Catholic News. 2008. “U.s. Birth Rate Spikes in 2006 Creating ‘Boomlet’” <https://www.catholicnewsagency.com/news/11476/us-birth-rate-spikes-in-2006-creating-boomlet>.
- Kearney, Melissa S., Phillip B. Levine, and Luke Pardue. 2022. “The Puzzle of Falling US Birth Rates Since the Great Recession.” *Journal of Economic Perspectives* 36 (1): 151–76. <https://doi.org/10.1257/jep.36.1.151>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- . 2021. *Tidyr: Tidy Messy Data*.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*.
- Wickham, Hadley, and Evan Miller. 2021. *Haven: Import and Export ‘SPSS’, ‘Stata’ and ‘SAS’ Files*.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.
- Yang, Yang, and S. Philip Morgan. 2003. “How Big Are Educational and Racial Fertility Differentials in the u.s.?*” *Biodemography and Social Biology* 50 (3-4): 167–87. <https://doi.org/10.1080/19485565.2003.9989070>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*.