Working Conditions in Amazon Warehouses and Per-Worker Productivity:

Exploring the Relationship Between Conditions for Full and Part-Time Warehouse Workers and Net Income Per Worker

Prachya Tantivatana
Master's Student of Quantitative Economics
University of California, Los Angeles

Tanner Woods
Master's Student of Quantitative Economics
University of California, Los Angeles

Jiacheng Zhu
Master's Student of Quantitative Economics
University of California, Los Angeles

Abstract — In this paper, we attempt to explore and estimate the cost of harmful working conditions in warehousing firms on worker productivity. More precisely, we attempt to implement financial and US workforce data from Amazon, Inc. alongside data on several industry averages to construct an approximate per worker valuation of productivity. For simplicity, this productivity is measured as a worker's share of Amazon's annual net income for a given monthly period; therefore we define Amazon's firm costs due to working conditions as a loss in productivity for a given worker. We use non-fatal workplace incidents (e.g. Injuries and illnesses) as a proxy for harmful working conditions. Additionally, we also control wages for working hours, trends, and time fixed-effects. Our models revealed few substantial significant outcomes, but did imply some degree of negative correlation. However, our results must ultimately be considered as inconclusive due to the compounding effect of layers of abstraction and a small sample size harming the robustness of estimation.

Keywords — Amazon, Amazon, Inc., deadweight loss, labor conditions, productivity loss, productivity per worker, working conditions

I. Introduction

In the world of labor conditions, Amazon, Inc. has a documented record of alleged labor violations internally and within partnered companies. Manufacturers sourcing parts for Amazon have been accused of slashing wages, relying upon underaged workers — which are officially termed as "interns" — and exceeding local working-hour limits[1]. In 2019, a Amazon warehouse in New Jersey failed to file reports on 26 work-related illnesses and injuries[2]. Former and current employees claimed that being "...vocally self-critical…" of one's shortcomings were an official staple of Amazon's leadership principles, while some claimed that the hustle-and-bustle of the Wall Street stock market floor paled in comparison to working at Amazon[3]. Meanwhile, departments secretly compete against one another to produce a product for greenlighting before their coworkers[4].

Despite these events, the arrival of an Amazon warehouse to an area often brings workers

[1] Day, Matt, and Debby Wu. "Amazon Under Fire Again as China Factory Hires Teen Interns." Bloomberg.com. Bloomberg, August 9, 2019.
[2] Soper, Spencer. "Amazon Faces Government Fine for Failing to Report Injuries." Bloomberg.com. Bloomberg, January 12, 2016.
[3] Kantor, Jodi, and David Streitfeld. "Inside Amazon: Wrestling Big Ideas in a Bruising Workplace." The New York Times. The New York Times Company, August 15, 2015.
[4] Ibid.

attracted by the promise of high wages relative to the locale. Furthermore, with a stated 2021 annual revenue of approximately 470 billion USD Amazon, Inc. exists as a titan within the business world. Despite the scandals, reports, and whistleblowers, it appears that nothing can halt the growth of the logistics-masked-as-ecommerce giant.

However, despite stunning growth and apparent success, are the working conditions of Amazon hindering its growth? Certainly, the company does not appear like it will depart from our modern world at any point soon. Yet, poor labor practices might ultimately be a prime example of Amazon cutting off its nose to spite its face: it wins in the short-term, but neglects possible worker productivity gains.

Unfortunately, there exists no company willing to air any alleged dirty laundry — existent or otherwise — for the world to examine. Amazon is no exception to this: detailed quantitative records of working conditions at Amazon or other companies with poor track records simply do not exist. Therefore, we will conduct our analysis on a synthesis of Amazon-specific and industry-wide data. Methodologies behind data collection and all extrapolations are detailed in **Section V —** for an abbreviated description — and **Appendix 1** — for a full breakdown. With the above in mind, it is critical to state that any results derived will only explore a general analysis of the relationship between working conditions and worker productivity. We cannot claim to possess data specific enough — and in enough quantity — to produce empirical claims on behalf of Amazon, Inc. .

In terms of overall layout, our paper is structured generally as follows. First, we will formally state the hypotheses of the authors. These are the primary questions that we will endeavor to produce answers for. Second, we will provide a short description of challenges encountered during the course of the paper. Third, we will explore the dataset gathered for analysis and any processing performed. Fourth, we will discuss the two primary regression models used within our analysis. Fifth, we will analyze the pertinent variants of the primary models. Finally, we provide our concluding statements on the outcomes.

First, we state our hypothesis and then discuss our methodology on data collection. Second, we explain our baseline model and its variation before analyzing model results in the fourth part. Finally, we evaluate this paper in the conclusion part.

II. Hypotheses

The authors of this paper will attempt to provide enough substantial and significant evidence to answer the following hypotheses:

Hypothesis 1: Workplace injuries/illnesses and days lost due to incidents have a negative relationship with worker productivity.

Hypothesis 2: Mean hourly wages have a positive relationship with worker productivity.

Hypothesis 3: Mean weekly hours of work has a positive relationship with worker productivity.

Hypothesis 4: Separation rate has a negative relationship with worker productivity.

III. Empirical Challenges

In the course of exploring the dataset backing this paper's analysis, we encountered only a single major obstruction: a simple lack of observations. However, this obstacle would prove to be a source of major concern from start to finish. Attempts to fit a regression produce a range of wildly different coefficients across model variants. We speculate that this is an attempt to fit onto the background noise of the model in lieu of more observational data.

Secondary issues arose in the form of "pseudo"-multicollinearity. Adjusted for inflation, wages have remained relatively stagnant — or even declined slightly. Incorporating the maintenance of an approximately 40-hour work week as the industry standard over the entire period of observation, we believe that regressions mistakenly interpreted the direction of stagnant wages and constant hours as evidence of multicollinearity.

IV. Data Description

For detailed documentation of model variables, please refer to **Appendix 2**; interaction variables are recorded in **Appendix 3**. Visualizations of year-wise and month-wise seasonality can also be located in **Figure 4**.

V. Data Processing

As mentioned previously, locating quantitative data on working conditions in Amazon proved to be challenging. Scant appropriate data was located; this data was furthermore strictly for the years of 2019-2021. As our team is attempting to track estimates of worker productivity in Amazon from 2011 onwards, this necessitated certain liberties surrounding the data. For a full breakdown of data processing, please refer again to **Appendix 1**. Additional information concerning transformations on non-linearly distributed covariates can be found in **Figure 2** and **Figure 3**. Furthermore, all data save reported workforce size, workforce composition, and quarterly net income is sourced from industry average series.

As a result of these methods, it is crucial to state that this dataset is not robust enough to provide more than an educated ballpark estimate of productivity relationships. Hopefully, a team with access to superior resources can acquire more precise data in any future study on this topic.

VI. Model Description

Our analyses are based on variants of the following primary model:

$$Y_t = \beta\, x_t + \delta V'_t + \gamma_t + \eta X'_t + \varepsilon$$

Where the regression terms are defined as:

- Y_i : Amazon's net income per worker in month t.
- x_t : Estimated monthly number of injuries and illnesses among Amazon warehouse workers.
- V'_t : Vector of controls including mean hourly wage, mean weekly hour work, separation rate, and trend of google searches for Amazon worker strikes[5] in month t.
- γ_t : Time fixed-effect including year and month.
- X'_t : Vector of interaction variables[6]
- ε : Error term

In total, each primary model received five fitted regressions. This paper will discuss the results of the most pertinent models, but more information can be located in **Table 1** for all models left uncovered.

[5] Intended as a means to capture workplace tension between Amazon and its employees. Can also represent the impact of increased public interest in events related to Amazon strikes.

[6] See **Appendix 3** for interaction variable components.

VII. Model Analysis

From the total body of fitted regression models, the most controlled model stands out in particular. This model shall be referenced as *M1-5*; referring to its position in column 5 of **Table 1**. For exact coefficient values of this model of interest, please refer again to the aforementioned location in the appendix.

Moving forward with our analysis, we shall begin with the first two models. In *M1-1* and *M1-2*, we are surprised to find that the results of the regression analysis revealed a statistically significant positive relationship between NFII and average income generated per worker, but others. Also, all other variables are insignificant. This result is counterintuitive and contrary to our hypothesis. We deduce that this "pseudo" positive relationship comes from the comovement of two sets of data: Amazon's income per worker is gradually increasing, while the average number of injuries per month is also increasing in the context of a significant increase in the number of employees. This can be a side note that Amazon has achieved a certain degree of co-growth in staff expansion and per capita efficiency. We believe that other variables, such as time trends, need to be controlled before we can better analyze the relationship between NFII and productivity.

In *M1-3*, we control the time fixed effect to get rid of the impact from the time trend. We can see that the results of the regression analysis are closer to our prior assumptions. Although the coefficient of injury and illness is not statistically significant, it starts to show a negative correlation. And the effects of median days miss and average weekly working hours on Amazon's average income are also significant at the 5% and 1% significance levels, respectively.

After the correlation analysis of the respective variables, we proceeded to introduce the interaction variables to better analyze the correlation between the variables (Please refer to **Appendix 3** and **Figure 1** for more details). In *M1-4*, injury and illness showed a negative correlation with work efficiency and was significant at the 1% level. However, the median days miss has a positive correlation with the dependent variable, which may also be caused by the time trend effect.

Ultimately, in *M1-5,* we control for both time fixed effects as well as interaction variables. Although the inflation of the coefficients themselves weakens the persuasive power of the model for quantitative analysis, the model presents a trend that strongly supports our prior hypothesis. More cases of injury and illness, as well as days of miss, would make firms less productive, while higher average hours worked and wages would provide a boost. Unfortunately, the coefficients of the separation rate were not significant in any of the four models, also preventing us from answering the question in hypothesis 4.

In addition to the non-constant term models, five additional constant models were constructed in the pursuit of thoroughness. These were abandoned in the face of hyperinflated constant values[7]. Furthermore, a constant should not be necessary for analysis in this situation; a situation where all covariates amount to zero — thereby producing a net income per worker of zero — is a valid potential observation.

VIII. Conclusion and future work

In recent years, workplace conditions such as safety are becoming a concern. Large-sized companies such as Amazon inc. are often criticized for labor exploitation through hazardous work environments to maximize company profit. It could be the case that firms do not take account of labor safety because they are maximizing profit. However, the principal-agent did show that

[7] For those maintaining an interest in the values of these constant-term models, please refer to **Table 2** within the appendix.

the profit maximizing firms have to take account of labor wellbeing so that they accept work contracts. In this paper, we quantitatively measured the impact of dangerous work conditions using injuries and illness on productivity. We found that the number of injuries and illness reported in the warehouse industry did negatively correlate with estimated Amazon inc. revenue per worker even after controlling for various factors. The result showed that there exist losses for firms from exploitation as expected.

Nevertheless, our results possess inflated coefficients, which we hypothesize is caused by an inefficient amount of our data and its quality caused by aggregating the data across different source and time frequencies with assumptions that might be unrealistic.

It is recommended for future work to improve upon the data aspect. For example, one may collect both productivity and workplace hazard using qualified measurement for multiple firms rather than just one to increase the robustness and generality of the result.

Appendix 1. Breakdown of all data processing methodologies and rationale

Variable	Components	Processing
Amazon Monthly Net Income	Amazon Quarterly Net Income	Equally split quarterly income over all respective three-month periods. As employee records implemented in this study do not include temporary or contracted workers, seasonal variations in per worker net income were deemed not of critical concern. Furthermore, any remaining seasonal trends might be dampened by quarterly reports covering a small enough slice of the year to capture some of the effect. Additionally, income is also inflation adjusted using the value of 2011 dollars.
Amazon Global Workforce	-	Monthly periods between annual reports in-filled as direct linear growth. Third-party reports of workforce size vary pre-2019 vary.
Proportion of Amazon US Workforce Laborer/Helper in Warehouse	-	Amazon, Inc. only began publishing EEO-1 reports beginning in 2019. Monthly periods between existent annual reports in-filled as direct linear growth. All dates pre-2019 are backfilled at a constant rate of ~38.98%; this is the proportion of US laborers/helpers to total global workforce in 2019. After 2019, Amazon, Inc. grew this proportion from ~38.98% to ~47.36% in 2021; this implies that applying the constant rate back to 2011 could be overly generous.
Industry Non-Fatal Injury or Illness Rate	Industry Non-Fatal Injury or Illness Rate	Equally split annual rate over all respective twelve-month periods.

(Monthly)	(Annual)	While this is the least robust extrapolation present, our team's rationale is that repeated injuries among full/part-time workers are uncommon and not subject to seasonal interference from inexperienced contractors/temps. If this assumption holds, then workers are about as likely to injure themselves at any point in the year.
Strike Google Trend		Monthly Google Trend Search Index for the keyword "Amazon worker strike" from 2011 to 2020. The data are obtained from Google Trend directly. Since it is an index, the data are ranging from 0 to 100 with 100 referring to the date with the most search volume within the search period. We incorporate this data because of potential heterogeneity for injury. To illustrate, some accidents in warehouses might be viral and cause the firm to lose more revenue due to publicity. As a result, we attempt to capture this "publicity" so that all injuries are treated equally.

Appendix 2. Definitions and documentation for baseline and time dummy model variables

Variable	Unit	Description
Amazon Net Income Per Worker	Billions, USD	Amazon's total net income is proportional to the percentage of total workforce composed of US warehouse workers, then divided equally among total US warehouse workers. Adjusted for inflation.
Amazon Non-Fatal Injury or Illness Estimate	log(Cases)	Industry rates of non-fatal injury or illness are set against Amazon's total US warehouse workforce to estimate Amazon's case numbers.
Non-Fatal Injury or Illness Rate	Percentage	Industry rates of non-fatal injury or illness.
Non-Fatal Injury or Illness Median Days Missed	Days	Industry monthly average of median days missed due to a case involving a non-fatal injury or illness.
Mean Weekly Hours	Hours	Industry monthly average for weekly hours worked for a warehouse worker.
Mean Hourly Wage	USD	Industry monthly average for hourly wage for a warehouse worker. Adjusted for inflation.
Separation Rate	Percentage	Industry monthly average rate for workforce separation.
Strike Google Trend	log(Index with respect to searches volume + 1)	Google Trend Search Index obtained from Google with the keyword "Amazon worker strike". The index corresponds to the relative search volume of the keyword from 2011 to 2020.
Year	-	Year (2011-2020)
Month	-	Month (January-December)

Appendix 3. Definitions and documentation for baseline and time dummy model variables

Interaction Component 1	Interaction Component 2
Mean Weekly Hours	*Mean Hourly Wage*
Mean Weekly Hours	*Separation Rate*
Strike Google Trend	*Separation Rate*
Amazon Non-Fatal Injury or Illness Estimate	*Non-Fatal Injury or Illness Median Days Missed*

Figure 1. Correlation Heatmap of Baseline Variables

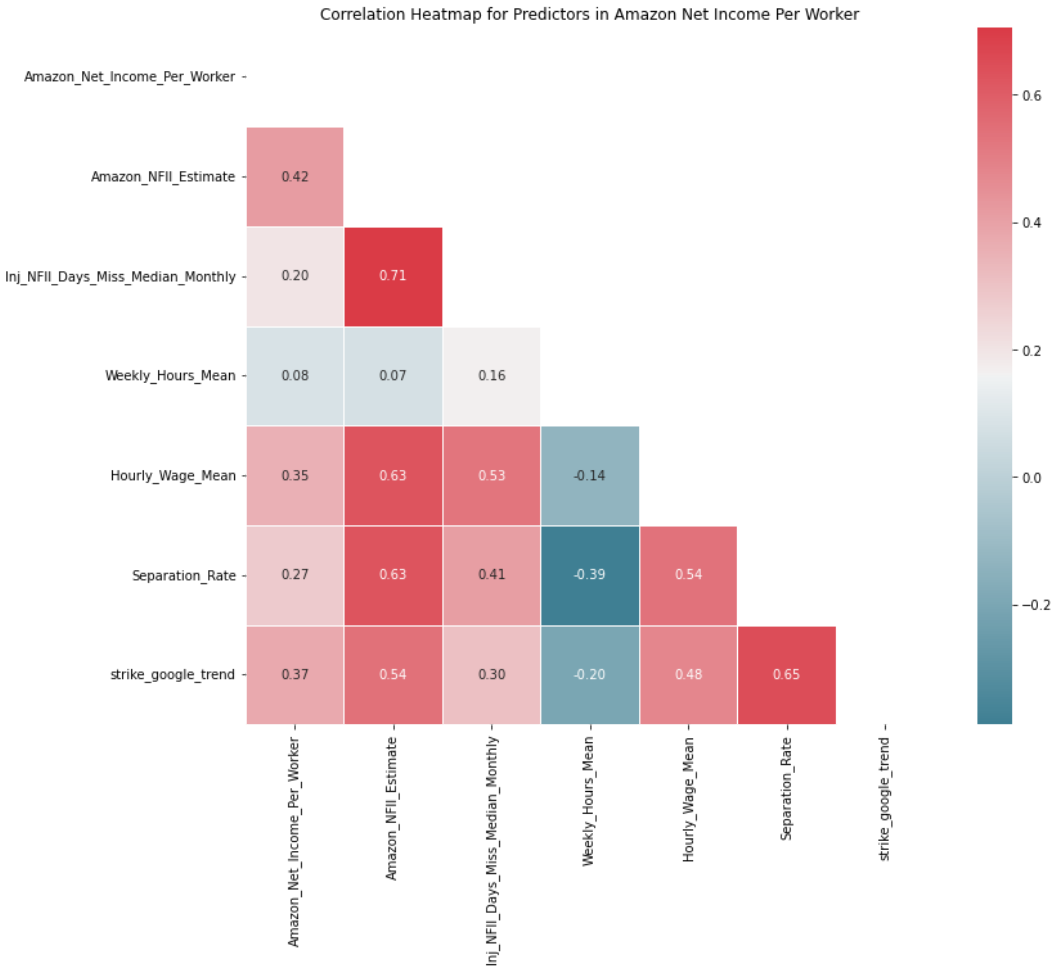

Figure 2. Quantile-Quantile Plots of Baseline Variables

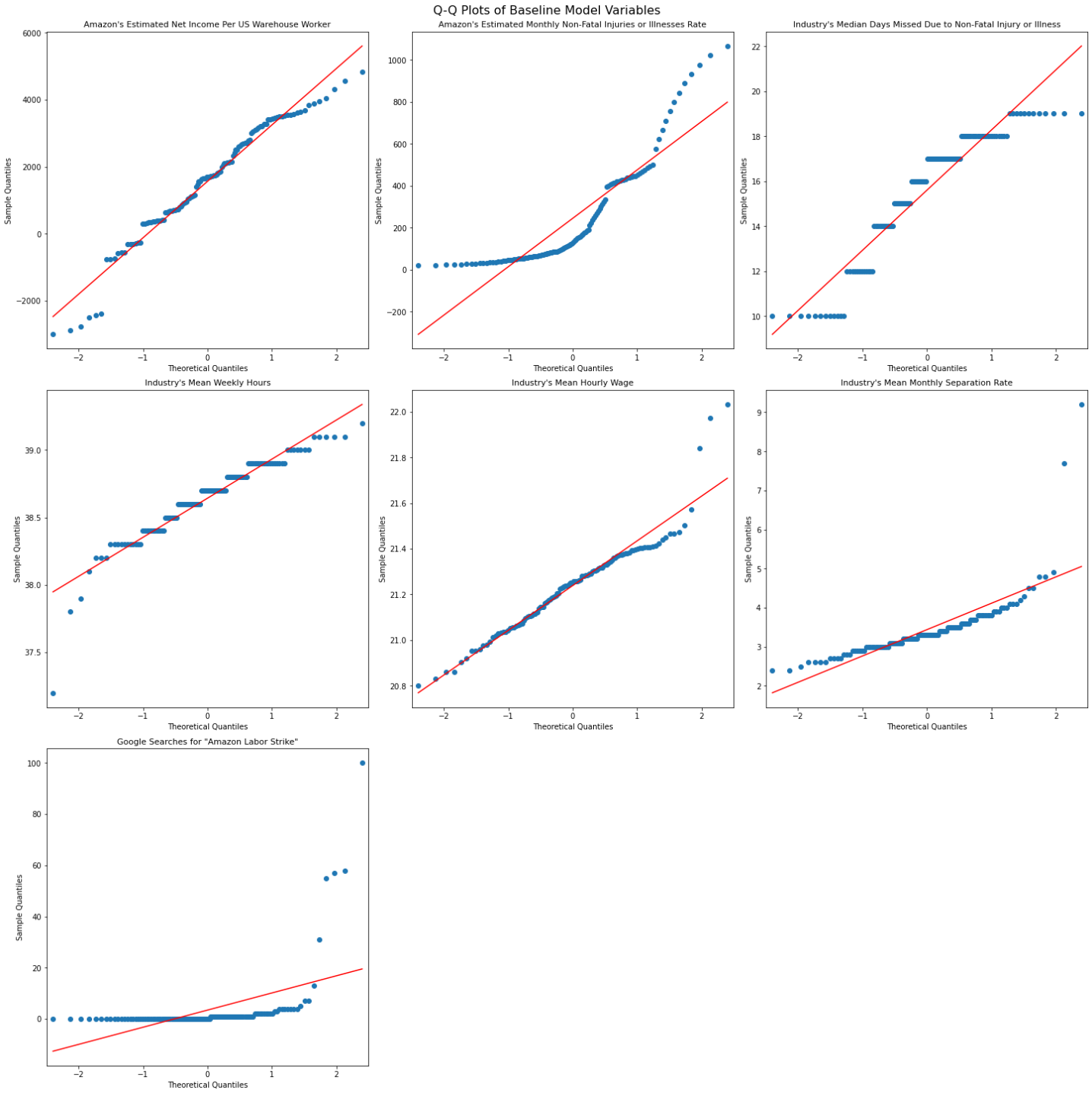

Figure 3. Comparative Quantile-Quantile Plots of Baseline Variables Receiving Transformations

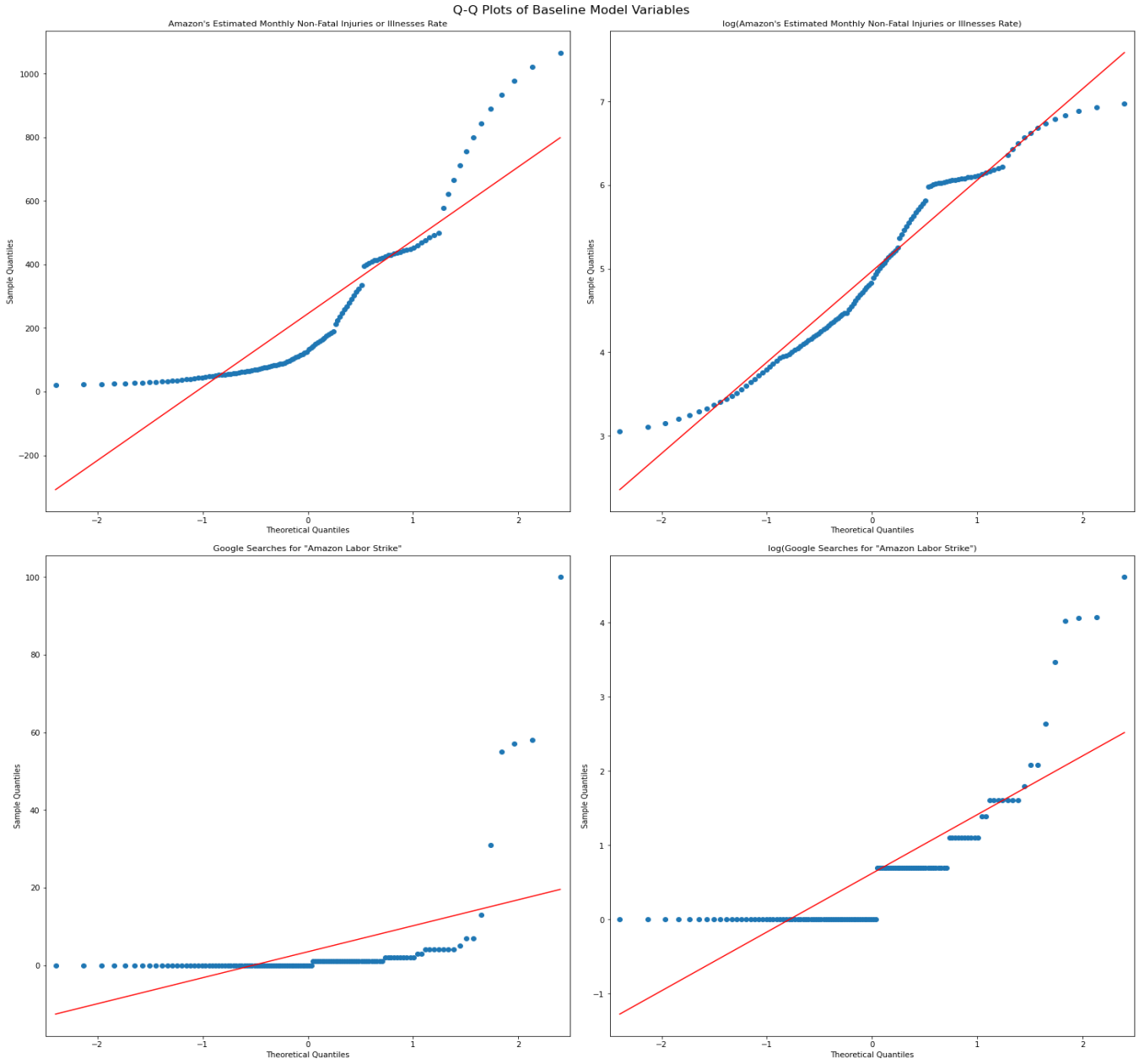

Figure 4. Year-wise and Month-wise Trends in *Amazon Net Income Per Worker*

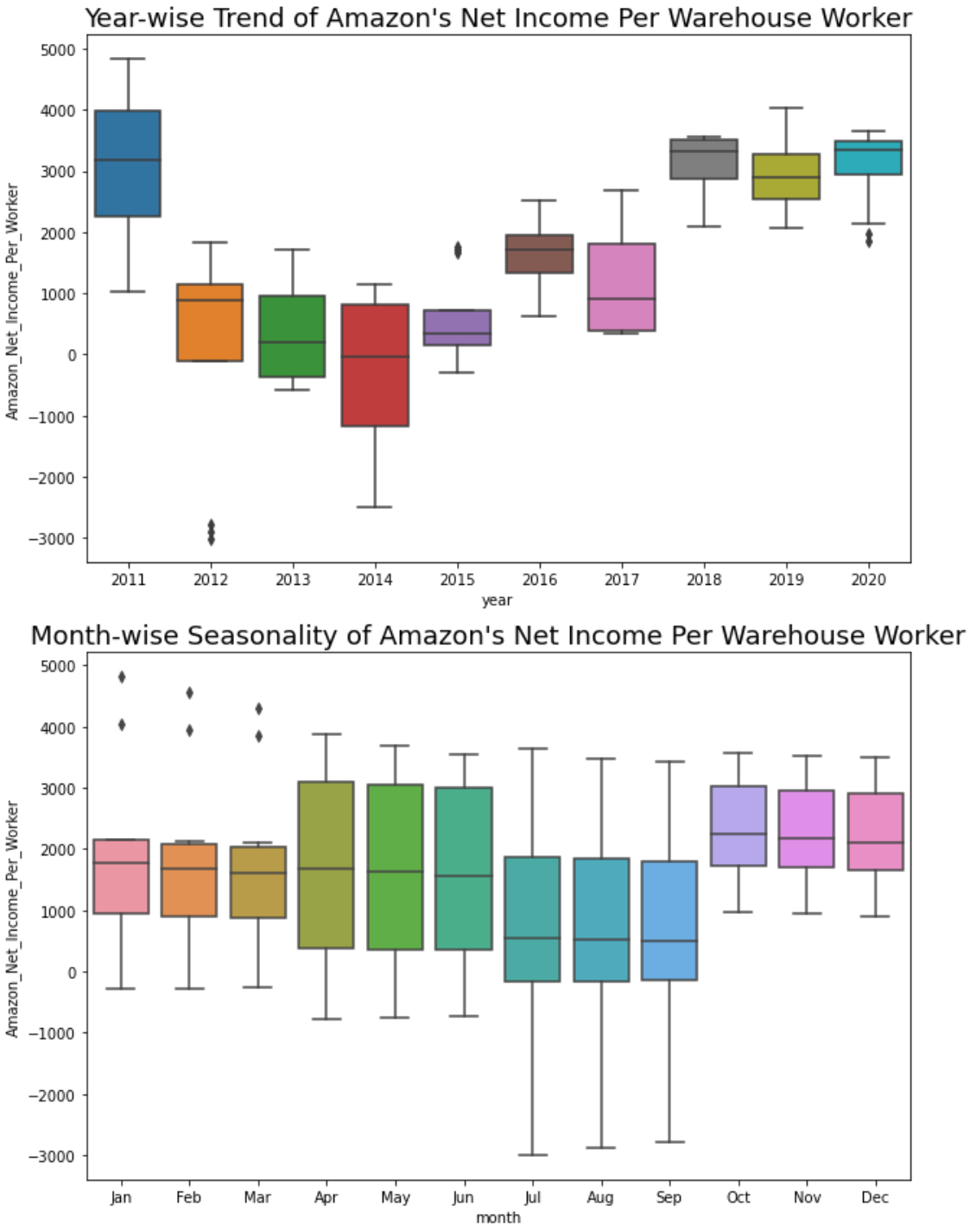

Table 1. Regression results of fitting non-constant term models

	(1)	(2)	(3)	(4)	(5)
Injury and illness	329.626 ***	710.918***	-1,877.689	-7,235.275***	-10,081.924**
	(27.840)	(218.240)	(2,045.114)	(1,164.332)	(4,320.601)
Median Day miss		-96.510	-5,515.911**	2,172.820***	-344,674.789***
		(71.790)	(2,132.438)	(306.511)	(98,411.162)
Average Weekly hour work		-87.728	1,619.366***	333.630	107,609.158***
		(351.935)	(464.951)	(481.200)	(30,650.719)
Average hourly Wage		165.478	618.669	-2,897.006	191,123.407***
		(661.219)	(755.467)	(3,325.190)	(57,270.275)
Separation Rate		-247.057	-13.164	19,251.846	-210.827
		(267.550)	(189.463)	(18,724.857)	(17,626.758)
Fixed Effect and Trend	No	No	Yes	No	Yes
Interaction	No	No	No	Yes	Yes
Observations	120	120	120	120	120
R^2	0.541	0.589	0.793	0.721	0.851
F Statistic	140.188	27.228	14.387	28.450	17.774

(1) Trimmed baseline; Amazon_Net_Income_Per_Worker ~ Amazon_NFII_Estimate

(2) Baseline; no year/month dummies or interactions

(3) Baseline + year/month dummies; no interactions

(4) Baseline + interactions; no year/month dummies

(5) Baseline + year/month dummies + interactions

Table 2. Regression results of fitting constant-term models[8]

	(1)	(2)	(3)	(4)	(5)
Const	1,620.098** (657.646)	-65,029.615** (31,512.195)	-41,932.373** (16,523.764)	-1,489,984.305 (1,616,100.359)	-2,415,365.416*** (694,366.086)
Injury and illness	641.028*** (129.314)	532.760** (231.858)	-1,877.689 (2,045.114)	-7,056.416*** (1,181.167)	-10,081.924** (4,320.601)
Median Day miss		-132.143* (72.862)	-2,021.546** (960.511)	-2,125.116*** (311.054)	-143,394.337*** (40,556.271)
Average Weekly hour work		843.674 (569.319)	1619.366*** (464.951)	38,900.611 (41,834.159)	107,609.158*** (30,650.719)
Average hourly Wage		1,572.861* (943.491)	618.669 (755.467)	70,015.944 (79,154.458)	191,123.407*** (57,270.275)
Separation Rate		-69.471 (277.488)	-13.164 (189.463)	691.997 (27,501.784)	-210.827 (17,626.758)
Fixed Effect and Trend	No	No	Yes	No	Yes
Interaction	No	No	No	Yes	Yes
Observations	120	120	120	120	120
R^2	0.172	0.249	0.793	0.476	0.851
F Statistic	24.573	6.254	14.387	9.886	17.774

(1) Trimmed baseline; Amazon_Net_Income_Per_Worker ~ Constant + Amazon_NFII_Estimate

(2) Baseline; no year/month dummies or interactions

(3) Baseline + year/month dummies; no interactions

(4) Baseline + interactions; no year/month dummies

(5) Baseline + year/month dummies + interactions

[8] As stated within the text, these models were abandoned for several reasons. All constant-term models above are included within the appendix as a matter of thoroughness, as well as to enhance our own prestige through an increased page-count.