The Effect of Franchise No-poaching Restrictions on Worker Earnings

Brian Callaci, Matthew Gibson, Sérgio Pinto,

Marshall Steinbaum, & Matt Walsh

October 2, 2024

RESTAT-29153: ReadMe File for Data and Code Disclosure

1 Introduction

This ReadMe file describes how to replicate the analysis and exhibits in "The Effect of Franchise No-poaching Restrictions on Worker Earnings." As the analysis in the paper proceeds in parallel with two sources, Burning Glass Technologies (BGT)[1] and Glassdoor (GD), there are two separate codeblocks described and indexed in this ReadMe file.

2 Obtaining the proprietary data

Since the paper uses proprietary data from BGT and Glassdoor, researchers wanting to replicate or extend our results will have to acquire the raw data from those vendors, respectively.

To inquire about BGT data, either contact the company using their online interface, available at https://lightcast.io/contact-us, or email Matt Walsh (matthew.walsh@lightcast.io).

To inquire about Glassdoor data, one should contact their Research team, listed at https://www.glassdoor.com/blog/research/meet-the-team/. The generic contact email is research@glassdoor.com. As of this writing (summer 2024), the two economists at Glassdoor are Daniel Zhao (daniel.zhao@glassdoor.com) and Aaron Terrazas (aaron.terrazas@glassdoor.com).

3 Software details

Glassdoor analysis was performed in Stata 17/MP running on Pop!OS (Linux) 22.04 LTS. BGT analysis was performed in Stata 17/MP running on Windows Server 2019 Standard edition.

[1] After we started working on this paper, BGT was involved in a merger and eventually changed its name to Lightcast. For simplicity, we keep referring to it as BGT, though the email contacts provided, for anyone who wishes to inquire about the data, reflect the newer designation.

4 File Directory

Links between GD output files and exhibits in the paper are specified in code comments within "Regressions_franchise_np.do."

Script	Description
Dataproc_GD.do	Imports CSV from GD, cleans, outputs as .dta
Dataproc_PCEPI.do	Imports CSV PCE deflator from FRED, outputs as .dta
PCEPI.csv	PCEPI deflator data from FRED
Dataproc_tightness.do	Computes quartiles of labor-market tightness in GD data
Dataproc_WA_AG.do	Imports CSV list of AODs and dates from WA AG, outputs as .dta
AOD_Filed_Date_20210308.xlsx	AOD dates provided by WA AG
Regressions_franchise_np.do	All GD figures and tables

Table 1. Glassdoor scripts and public data

Script	Description	Exhibits generated
A1 - deflator pcepi.do	Imports FRED's csv PCE deflator file, outputs as .dta	
A2 - usda county-to-cz crosswalk.do	Imports USDA's csv county-to-CZ crosswalk, outputs as .dta	
A3 - wa ag list of chains and settlement dates.do	Imports WA AG's list of chains and settlement dates in csv, outputs as .dta and .xls table	Table A.1
A4 - franchise_name-franchise_id correspondence.do	Imports the initial list of 530 franchise names and numerical identifiers, and adds the franchise chains that were on the WA AG list but not in the original set of 530. The chains that did not sign AODs from the list of 530 constitute our placebo treatment group, on which event studies are conducted in script 3c.	
A5 - naics-4d download and cleaning.do	Downloads, cleans, and converts to .dta the 2017 NAICS 4-digit codes and names	
1a - bgt year-month files + year files + append full 2007-2021.do	Unzips and extracts all year-month txt files, converts to .dta, and appends them to create a complete 2007-2021 file	
1b1 - tag franchises in BGT + merge with CZ.do	Tags franchises in the BGT data	

1b2 - loop all year-month + create full BGT file tagged with franchises.do	Creates each year-month BGT file with the unique franchise identifiers and appends them for the entire time period	
1c - baseline file for regressions.do	Creates the base files to be used in implementing the empirical specification	
2a - list of treated industries.do	Creates a file with the list of treated industries	
2b - list of restaurant chains.do	Creates a file with the list of restaurant franchise chains that were treated	
2c1 - laus + bgt - labor market tightness indicator.do	converts and appends the monthly LAUS files and creates a labor market tightness indicator for 2017m1-2021m12, at the CZ level	
2c2 - chain-level tightness indicator and quartiles.do	Creates a file with the average tightness faced by each chain and the tightness quartile they are assigned to	
3a1 - summary stats - tables a4 and a5.do	Computes summary statistics tables	Tables A.4 and A.5, columns 3-4
3a2 - summary stats - top industries.do	Computes the number of observations by 4-digit NAICS, for the same-industry sample and the inverse control group	Table A.3
3a3 - summary stats - same-industry - obs with and without estimates.do	Computes the number of observations, with and without Indeed/LinkedIn ads, same-industry sample	Figure B.1(A)
3a4 - summary stats - inverse - obs with and without estimates.do	Computes the number of observations, with and without Indeed/LinkedIn ads, inverse sample	Figure B.1(B)
3b1 - regression tables and figures - same + inverse - baseline.do	Estimate event study specifications in the same-industry sample & generate associated plots	Figure 1(A), 1(C), B.2(A), and B.2(C); Table 1 (columns 1 and 3)
3b2 - regression tables and figures - same + inverse - no indeed and linkedin.do	Implements the specifications that exclude the job ads coming from Indeed and LinkedIn	Figures B.2(B) and B.2(D); Table B.1
3b3 - regression tables and figures - same + inverse - pre-March-2020.do	Implements the specifications that exclude job ads from the post-March 2020 period	Figures C.1(A) and C.1(C)
3b4 - regression tables and figures - same + inverse - by lm_tightness.do	implements the specifications that compute the effects by labor market tightness quartile	Figure D.1(A) and D.1(C)

3c - regression tables and figures - same + inverse - non-AOD chains.do	Implements a robustness check where a placebo test is conducted, using franchise chains that did not enter into an AOD	Figures 2(A) and 2(C)
3d1 - regression tables and figures - same - by pay frequency.do	Implements the specifications that split the sample between hourly wage and annual salary workers, using the same-industry sample	Figures A.1(A) and A.1(C)
3d2 - regression tables and figures - inverse - by pay frequency.do	Implements the specifications that split the sample between hourly wage and annual salary workers, using the inverse sample	Figures 3(A) and 3(C)
3e - regression tables and figures - same + inverse - 6d occupations.do	Implements the specifications that use 6-digit occupations (instead of 4-digit occupations)	Figures A.2(A) and A.2(C)
3f - regression tables and figures - inverse - restaurants.do	Implements the specification that excludes non-restaurant chains from the treatment group	Figure E.1(A)

Table 2. BGT scripts and the exhibits they generate.

5 Data dictionaries

Table 3 describes native variables (fields) in the analysis data set "GDsalaries_cleaned.dta." To understand variables computed by us, see our GD scripts, particularly "Dataproc_GD.do." GD fields are as named in the CSV file provided to us. We have observed that GD sometimes renames fields when pulling a flat file from their relational database, so future researchers may see different names on the same variables. We have also observed that GD sometimes changes the construction of a variable while preserving the same or similar name. For most identifier variables, the data processing script "Dataproc_GD.do" applies corresponding string fields as value labels and discards the string fields to reduce the memory footprint of the data set. Exceptions to this rule are employerid and jobtitleid, which take on too many values to permit the use of value labels, and metroid.

GD field	Description
reviewid	Unique report identifier
employerid	Employer identifier
userid	User identified
jobtitleid	Job title identifier
jobtitle	Job title (string)
basepayamount	Base pay (may be hourly, annual, or monthly)
cashbonusyearly	Cash bonus (annual)
stockbonusyearly	Stock bonus (annual)
profitsharingyearly	Profit sharing (annual)
salescommissionyearly	Sales commission (annual)
tipsyearly	Tips (annual)
yearofsalary	Calendar year of salary (may differ from report year)
yearsrelevantexperience	Years experience (many users appear to interpret as job tenure)
metroid	MSA identifier
metroname	MSA (string)
birthyear	Birth year
employeestotalnum	Firm headcount
employername	Employer name (string)
occupationgroup	Specific occupation identifier
majoroccupationgroup	General occupation identifier
genderid	Gender identifier
employmentstatusid	Employment status identifier (e.g. full-time, part-time, contract)
basepayperiodid	Period of base pay (hourly, monthly, or annual)
contentoriginhookid	How user came to the GD site
stateid	US state identifier
highesteducationid	Educational attainment identifier
industryid	Industry identifier
sectorid	Sector identifier
employertypeid	Employer type identifier (e.g. public company)

Table 3. Glassdoor data dictionary

For BGT, the following table identifies the variables included in the data used for this paper, which was obtained in April 2023. We have observed that different vintages of BGT data may include different variable names for the same underlying variable, and more significantly, includes additional fields as time goes on. The fields named in the following table are those from the April 2023 vintage data used for the present project.

BGT field	Description
bgtjobid	Unique ID generated by BGT
employer	Employer name
soc	SOC code of the job, assigned using BGT occupation coding rules, based on the 2010 SOC delineations
naics4	4-digit NAICS code
fips	5-digit code, representing the concatenation of the state and county FIPS codes
minsalary	Minimum annual salary. If hourly rate is provided, this rate is converted to annual salary through multiplying the rate by 2080
maxsalary	Maximum annual salary. If hourly rate is provided, this rate is converted to annual salary through multiplying the rate by 2080
jobdate	Date the posting was spidered
Researcher-created field	
franchise_id	ID that uniquely identifies franchise chains
franchise	Franchise chain name, also uniquely identifies franchise chains
naics4_contracts/naics4_ag	Franchise chain-level NAICS 4-digit industry, sourced from either FDDs or the WA AG's AODs

Table 4. BGT data dictionary