THE UNIVERSITY OF CHICAGO

ESSAYS IN LABOR ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
NING DING

CHICAGO, ILLINOIS
JUNE 2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply grateful to my advisors, Magne Mogstad, Stéphane Bonhomme, and Thibaut Lamadon. Chapter 1 is my job market paper. I thank Lasse Eika for helpful comments and extremely valuable support for data. I am also grateful to Kuan-Ming Chen, Jake Torcasso, and seminar participants at Econometrics working group and Professor Magne Mogstad's research group at the University of Chicago for detailed comments. I thank Statistics Norway for data provision.

Chapter 2 is joint work with Kuan-ming Chen, John List, and Magne Mogstad. The opinions expressed in the chapter are those of the authors alone and do not reflect the views of Uber. We appreciate the constructive comments and suggestions from Oystein Daljord, Jonathan Hall, Vishal Kamat, Libby Mishkin, Jesse Shapiro, Winnie van Dijk, and Kerry Smith.

The past seven years presented gifts as well as challenges. I would like to additionally thank Leah Luben, Yu-ting Chiang, Mauricio Chikitani, and Maxwell Kellogg for sharing both joy and tears together. These people made Chicago home to me. I also want to thank Liyun Ju, Ophelia Fang, Cindy Tian, and Karen Ju who have supported me throughout the program, even though we don't get to reunite often.

I owe a huge debt of gratitude to my family who has unconditionally supported me to pursue my ambition and made sure that I know they will always be my home. Lastly, I want to remember my late grandmother who passed away 4 years ago. Her last words will continue to be a source from which I draw strength.

To my parents.

ABSTRACT

This dissertation consists of two parts. The first part concerns how and to what extent the worker-firm rent-sharing mechanism confounds the estimates of the return to tenure. This study finds that the worker-firm rent-sharing responses empirically create a large downward bias in the return to job tenures in the context of the Norwegian labor market. I conduct detailed analyses on potential sources of biases. I argue that workers' job mobility not only depends on the differences in the origin and destination firms' time-invariant pay premiums but also on the realizations of their firm-specific wage innovations. I show that the between-firm differences in the firm-specific wage innovations play a dominating role in driving the large downward bias in the previous return-to-tenure estimates. The estimated firm-specific time-varying pay premiums are positively correlated with but not sufficiently explained by firm-level productivity measures such as value-added or firm sizes. Instead, I find that alternative measures, such as employment shares of newly hired workers, which capture the booms and busts and the compositional changes in employment at the firm level, are good proxies to help reduce biases.

In the second part, we quantify how labor supply elasticities and reservation wages vary between people and over time, and infer workers' valuation of flexibility in their choices of work hours. Economists and policymakers are keenly interested in these quantities, especially lately with the growth in jobs that offer flexible work schedules. Our study takes advantage of a large natural field experiment at Uber, the largest ride-sharing company. Combining this experiment with high frequency panel data on wages and individual work decisions, we estimate a dynamic labor supply model that let us recover reservation wages, labor supply elasticities, and workers valuation of flexibility.

CHAPTER 1

BIASES IN RETURN TO TENURE, JOB MOBILITY AND WORKER-FIRM RENT-SHARING MECHANISMS

1.1 Introduction

Whether and by how much wages rise with job tenures in a worker-firm employment relationship have received heated debates in the literature. Assessing the magnitudes of the return to job tenures provides key implications for understanding how much a worker's wage growth is tied to employers and what explains the shape of life-cycle earnings of workers with different characteristics. Yet, quantifying the magnitude of the return to job tenures has faced challenges regarding potential selection biases and the measurement of wages and job tenures. Although wages rise extensively with job tenures in the cross-section ([48]), previous work, such as [2] and [3], have pointed out that the cross-sectional relationship between wages and tenures may reflect selection biases rather than a positive return to job tenure. If worker-firm matches differ in their pay levels, and if an employment relationship tends to last longer for higher-paying job matches, then a positive relationship between wages and job tenures can emerge in the cross-section, even though a worker's wage does not rise as he works one more year on his job. Namely, the return to tenure estimates based on the cross-sectional data suffer from an upward bias from a job match perspective.

While previous literature has developed to address the unobserved factors that determine wages and affect job mobility, they typically focus only on time-invariant unobserved components (except for [55]). Such components include worker and match fixed effects (see, for example, [56], and [4]). However, the growing literature on the worker-firm rent-sharing mechanism in imperfectly competitive labor markets has brought forth evidence that within-firm changes in workers' wages partially reflect firm-level shocks (for example, [28], [43], and [27]). The implied variation in the firm pay level across time gives rise to a rarely discussed potential source of bias in the return to job tenure.

This paper considers identifying and estimating the return to job tenure in a labor market environment where firm-specific pay policies may differ across firms and fluctuate over time. I adapt a Mincerian-style wage model and combine it with [1]'s framework, which allows for flexible sorting patterns based on worker and firm heterogeneity. Importantly, I consider firm unobserved heterogeneity along two dimensions: the time-invariant heterogeneity which reflects differences in firms' permanent pay policies, and the time-varying heterogeneity that captures the common wage fluctuations within firms. I argue that biases in the return to job tenure estimators arise because selection on job moves is not only based on origin and destination firm's fixed pay premiums, but also on within-firm and between-firm differences in the firm pay innovations. The paper finds that the time-varying firm heterogeneity introduces a downward bias in the estimates of the return to tenure. In addition, the estimated return to tenure is mostly concentrated during the first few years of a job in the Norwegian labor market.

I begin by describing the data in Section 1.2. This study uses the Norwegian matched employer-employee data which spans over 20 years and covers the universe of workers and firms in Norway. The long panel allows me to track workers' job histories in detail, including firm identities, industry experiences, and general labor market experiences. Moreover, the rich records on firm balance sheets permit detailed investigation of the relationship of the time-varying firm pay policies with external measures of firm-level outcomes, including value-added, operative revenues, and total sales. Another advantage of the Norwegian administrative data is that its rich information on observed worker characteristics allows me to study the sources of wage growth for workers with different demographic and social economic backgrounds.

The formulation of the wage model in this study follows the literature. Two variables of interest compete in driving a worker's wage growth: the worker's tenure at his current firm and his total labor market experiences. I emphasize three types of unobserved heterogeneity: worker unobserved productivity of earning wages, differences in firms' pay premiums which are fixed over time, and a time-varying component in firms' pay premiums that differ across firms and time but are common to co-workers within firms. To account for the time-varying feature of firm pay premiums, I include a set of fully interacted firm×year fixed effects in the wage model, in addition to worker fixed effects to capture the heterogeneity in workers' productivity. Thus, this framework resembles that of [1] two-way fixed effects model and offers a transparent setup where co-workers' wages are allowed to co-move over time, without explicitly specifying the rent-sharing mechanism. The identification of the return to job tenure in this wage model invokes a set of assumptions similar to, but less restrictive than, the exogenous mobility assumptions in AKM. In particular, it allows worker mobility and sorting based on not only worker and firm permanent heterogeneity but also on the innovations in wages at the firm level. Allowing mobility to depend on the firm-level wage innovations is crucial to correcting biases in the return to tenure.

Motivated by several pieces of evidence on the importance of worker and firm time-invariant heterogeneity and the co-movement between workers' wages and firms' productivity, I estimate the wage model with time-varying firm fixed effects. There are three main findings.

First, I find a substantial downward bias in the return to job tenure estimates due to the omission of time-varying firm heterogeneity. For male workers with college degrees or above, the estimated total return to tenures in the first two years of an employment relationship is 8.7 log points, about a 13% increase from the estimates based on alternative wage specifications. My estimates are comparable to the estimates obtained using [56]'s approach, and they are much larger than what [3]'s strategy suggests. I argue that the downward bias arises because workers' job mobility is a function of not only the origin and the destination firms' fixed pay levels, but also the realizations of these firms' time-varying pay shocks. Two forces are at play. On the one hand, the relationship between a mover's foregone tenure levels and the within-firm changes in pay levels at the origin firm contributes to an upward bias in the return to tenure. An exceptionally low growth in the firm's pay level from one year to another induces a worker to quit early to avoid absorbing the negative wage growth.[1]

1. In this study, I do not distinguish between layoffs and quits. However, the differences between voluntary and involuntary job separations are important in understanding job mobility and tenures.

All else being equal, a more senior worker is willing to quit and give up the accumulated tenure when there is a more negative shock to the firm's pay levels. As a result, high movers' tenures are associated with low (origin) firm-level pay innovations, contributing to an upward bias. On the other hand, the relationship between a mover's foregone job tenures and the differences in the time-varying firm pay innovations between the origin and the destination firms creates room for a downward bias in the return to tenure. A mover with a higher tenure will require a higher firm pay innovation at the destination firm to compensate for his loss of the previously accumulated tenure. Hence, the newly reset movers' tenure is associated with high (destination) firms' pay innovations, leading to a downward bias. Therefore, these two competing forces together drive co-movement between movers' tenure levels and firm pay innovations. Empirically, the between-firm growth effect dominates, leading to a net negative bias in the estimates based on alternative model specifications. I verify how these forces generate biases in the Monte Carlo simulation exercises in Section 1.7.

The second finding is that, while the return to tenure is substantial in the first two years at a job, it falls back to almost zero after the first two years at a job in the Norwegian context. My estimates reveal that male workers with college degrees or above enjoy a total wage return to experiences of around 8.6 log points in the first two years into the labor market, similar to the return to tenure (8.7 log points). Different from the return to tenure, the general labor market experiences remain a driving force for wage growth in the first ten years of these workers' careers. The estimated return to ten years of experiences is about 25.4 log points.

The third finding relates the estimated time-varying effects to firm-level performance measures. Taking advantage of the rich information from the firm balance sheets, I find that the firm-level productivity measures, such as value-added and firm sizes, have a high explanatory power for the variation in the level of the estimated time-varying firm fixed effects in the cross-section. The estimated elasticity between value-added and the estimated firm fixed effects is around 0.02. However, the explanatory powers of the firm productivity measures drop substantially when within-firm changes are examined. Moreover, when being used as proxies for the time-varying firm pay components in the estimation of the wage model, I find that these firm productivity measures eliminate the downward bias insufficiently. Instead, I find that measures such as firm employment shares of newly hired workers perform well in reducing the bias in the return to tenure estimates. Such variables can be viewed as capturing both the booms and the busts in employment at the firm level, accounting for factors that attract workers and raise wages other than productivity shocks (e.g., amenity shocks).

Lastly, in Section 1.8, I distinguish between and assess the relative magnitudes of the returns to firm and industry tenures, among male and female workers with different educational backgrounds. I find that the return to firm tenures diminishes but does not entirely disappear once industry-specific capitals are accounted for. Overall, the returns to job tenures and experiences serve as the primary source of wage growth in the first few years after workers' entrance into the labor market. The preferred estimates of the return to job tenures are comparable for workers from the same gender group. The comparison along the gender dimension reveals that the female workers receive around half of the male workers' return in the first 2 years of employment. Compared across education groups, the return to the general labor market experiences is substantially larger and remains to be large over time for workers with higher education degrees. These workers enjoy an average of 1.7% growth in wages per year in the first ten years. For workers with high school degrees, wage growth from the accumulation of general labor market experiences is substantial (roughly 1.1% per year) in the first five years. However, it falls back to almost 0 afterward. Industry returns remain small but non-trivial across all groups of workers.

This study relates to the extensive literature that studies the reduced-form estimation of the causal effect of job tenures on wages. This strand of literature is concerned with correcting the bias due to the presence of unobserved heterogeneity that affects wages and job mobility. Typical unobserved factors of interest include worker productivity and worker-firm complementarities, both of which are fixed within job spells (see [2], [3], [56], and [4]). The closest work to this study is [55]. They point out that the variation in the firm pay policies over time creates a downward bias in the return to tenure estimates using Portugal and Germany data. They conjecture that this bias is related to that firms increase hiring when they receive better shocks, and thus including firm employment sizes as control variables corrects the bias to a large extent. This study complements but extends their work in several important ways. Different from their work, I embed an AKM style framework into a Mincerian wage model. With this framework, I invoke a different set of identifying assumptions with which I provide detailed analyses and straightforward intuitions of the biases stemming from various sources. Moreover, I directly show that the time-varying firm pay components are positively related to but not sufficiently explained by firm productivity measures such as log value-added or firm employment sizes. Firm-level variables such as the employment share of newly hired workers seem to be good proxies for the time-varying firm pay components and help eliminate the bias in the return to tenure estimates.

The second strand of literature studies the return to industry tenure (for example, [49], [50], and [22]). [49] is among the first to empirically assess the importance of the return to industry tenure. He focuses on worker displacements following establishment closures. By examining the wage loss between those who switch industries and those who stay in the same industries, he finds that the wage cost associated with the displacement is more considerable for industry switchers, especially those with pre-displacement industry tenures. He interprets these results as evidence of the presence of industry-specific skills. [22] build a framework where a worker's wage is modeled as a function of firm tenures, industry tenures, and general labor market experiences, allowing for rich heterogeneity in worker and firm pay profiles. They rely on displaced workers from firm closures as well in order to mitigate the selection biases (for example, inter-industry job mobility may occur because workers find improved matches in another industry). This paper studies the relative magnitudes of firm tenures, industry tenures and experiences by accounting for time-varying firm and industry effects and invoking the exogenous mobility assumption for identification.

The study also contributes to the growing literature on worker-firm rent-sharing mechanisms. This literature studies the pass-through of the fluctuations in firm-level productivity shocks to incumbent workers' outcomes (see, for example, [57]; [28], [13], [40], [23], [41], [6], [27] and [43]). I complement this literature by first showing how worker mobility/job turnovers and firm-level employment composition change with respect to firm-level productivity shocks, particularly concerning the average worker tenures. I adopt a similar framework to [23] and [41] where I extend the traditional AKM two-way fixed effects model by introducing firm×year effects in workers' wages. This framework offers a transparent setup where co-workers' wages are allowed to co-move over time, without specifying the pass-through mechanism explicitly. I also build the bridge between this literature and the large literature on the return to firm tenure. I show that the rent-sharing mechanism plays a key role in quantifying the return to job tenures.

The paper is organized as follows. In Section 1.2, I describe the data sources. Section 1.3 introduces the wage model and the identification assumptions. In Section 1.4, I provide analyses on the bias in the return to tenure estimates resulting from worker and firm permanent unobserved heterogeneity. Section 1.5 motivates and shows the estimation results of the return to job tenure when firm time-varying pay components are taken into account. Section 1.6 derives biases and explores the use of external measures of firm-level outcomes as controls. In Section 1.7, I conduct Monte Carlo simulation exercises to verify the intuition of biases stemming from various wage specifications considered in the paper. Lastly, in Section 1.8, I compare the wage returns to firm tenures, industry tenures, and general experiences across workers of different observed characteristics. Section 1.9 concludes.

1.2 Data

The data source. This study uses data from the State Register of Employers and Employees in Norway, covering the universe of firms and workers who resided in Norway for at least six months between 1995 and 2018. For each job spell, detailed information on the number of days worked per year, the contracted hours per week, and the annual earnings are recorded.

Given that the Norwegian data provides contracted hours per week and the number of days worked each year, I take advantage of this detailed job spell information and construct average hourly wages as the wage measure for this study.

Sample selection and variable definitions. This study restricts to the workers aged between 18 and 64 who have at least one job spell in the private sector in the sample period. I drop self-employment spells and further restrict to the workers who work full-time (contracted hours $\geq$ 30 hours/week) and full-year (days worked $\geq$ 180 days per year) in each calendar year.[2] For each observed job spell, I drop the last year to exclude end-of-job irregular payments such as bonuses and severance pay. Lastly, I drop job spells that are observed to start in 1995 to avoid left censoring. Imposing the above sample criteria results in a sample consisting of 9,661,253 total observations, 1,273,384 unique workers, 215,400 unique firms, and 2,381,375 unique job spells.

To allow for heterogeneity in the return to job tenures across men and women, skilled and unskilled, the estimation sample is further divided into 4 groups based on a worker's gender and education level. Workers fall into 2 education groups: those with high school degrees, and those with college-equivalent degrees or above. Table 1.1 and Table 1.2 report the summary statistics on job records and firm-level measures. In the appendix, I show the distribution of tenure across the four groups of workers.

2. In the case where a worker works for multiple firms in a year, I select the highest-paying firm for the worker.

	All	College Male	HS Male	College Female	HS Female
Mean log hourly wage	3.6799	3.9196	3.6507	3.6572	3.3381
SD of log hourly wage	0.4777	0.4798	0.4257	0.4477	0.4477
Mean tenure	3.6204	3.5714	3.9533	3.30794	3.3185
Mean experiences	8,3806	9.0639	8.6081	7.1517	7.3286
Age	38.7051	40.5204	38.1319	38.0323	37.9967
Mean spell length	7.1204	7.1162	7.6874	6.1190	6.5569
Married	0.4312	0.5095	0.4000	0.4300	0.3892

Table 1.1: Summary Statistics

	Mean	SD
Firm log value-added	13.7102	2.0034
Firm log sales	14.7690	2.2305
Firm log operative revenue	11.8493	2.3298
Firm full-time employment size	4.4243	20.6602
Firm age	6.9194	5.1093

Table 1.2: Summary Statistics of Firm Characteristics

Notes:
[1] The summary statistics are obtained using a sample of unique firm × year pairs. Therefore, the estimates are not weighted by employment sizes.
[2] To avoid left censoring in measuring firms' ages, I report the mean and the standard deviation of firm ages on a sub-sample of firms where firms are first observed and recorded strictly after 1995.

1.3 The Wage Model

I consider the following wage model for worker i in year t:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \gamma X_{it} + \varepsilon_{ij(i,t)t} \tag{1.1}$$

where w_{it} denotes the log real wage of worker i in year t, $T_{ij(i,t)t}$ measures worker i's job tenure with firm $j(i,t)$ at the beginning of year t, M_{it} measures the actual years of labor market experiences at the start of year t, and X_{it} is a vector of observed worker characteristics.[3] $j(i,t)$ is a function that indicates the employer of worker i in year t. In

3. Worker-level characteristics include the marital status and occupation.

the estimation, I allow $T_{ij(i,t)t}$ to be fully flexible in the estimation by including a series of dummies for different tenure levels. The labor market experiences, M_{it}, enter the wage equation in the form of a 6th-degree polynomial.

The unobserved term $\varepsilon_{ij(i,t)t}$ consists of the following components:

$$\varepsilon_{ij(i,t)t} = \alpha_i + \psi_{j(i,t)t} + \phi_{ij(i,t)} + u_{it} \tag{1.2}$$

where α_i measures a worker's permanent ability to earn wages that is portable across firms, $\phi_{ij(i,t)}$ captures the complementarities between worker i and firm $j(i,t)$, and u_{it} is worker-level idiosyncratic wage innovation. As a key addition to the wage model, $\psi_{j(i,t)t}$ captures the heterogeneity in firms' pay premiums that vary over calendar years. This time-varying firm component introduces a common wage shock to all incumbent workers employed at the same firm in the same year, capturing the worker-firm rent-sharing mechanism. Without loss of generality, I decompose $\psi_{j(i,t)t}$ into three components:

$$\psi_{j(i,t)t} = \eta_{j(i,t)} + \tilde{\psi}_{j(i,t),t} + \delta_{r(i,t),t} \tag{1.3}$$

where $r(i,t)$ is a function that maps worker i in year t to his corresponding local labor market. $\delta_{r,t} = E[\psi_{j(i,t)t} | r(i,t) = r, t]$ is the market×year fixed effect that captures the aggregate shocks to wages that are common to all workers in the same local labor market. A local labor market in this study is defined as an industry×labor market region pair in Norway.[4] $\eta_{j(i,t)}$ is the firm time-invariant fixed effect which can be interpreted as the firm-specific average pay premium that is fixed over time. Firms may differ in their time-invariant pay premiums as a result of the differences in firms' permanent productivity. I assume that $\eta_{j(i,t)=k} \equiv \frac{1}{T_k}\sum_{t=1}^{T_k}(\psi_{kt} - \delta_{r(k,t),t})$, where T_k is the observed number of periods of

4. Firms are collapsed into ten industry groups: manufacturing, utilities and transport, construction, trade, services, business, education, health, public administration and other.

firm k in the panel. Lastly, $\tilde{\psi}_{j(i,t),t}$ is the time-varying firm fixed effects that reflect the idiosyncratic fluctuations in a firm's pay policies around the mean. Due to the inclusion of the market×year fixed effects, and the firm time-invariant fixed effects, $\tilde{\psi}_{j(i,t),t}$ is mean zero for each firm, i.e., $\frac{1}{T_k}\sum_{t=1}^{T_k}\tilde{\psi}_{kt} = 0$ for all k.

The key parameters of interest in this study are β_1 and β_2. β_1 is the causal effect on wages of a worker spending one additional year with his current employer, holding the total labor market experiences, the worker's permanent ability of earning wages, the firm's (time-varying and time-invariant) pay policies, aggregate market shocks, and the worker-firm complementarity effect constant. I refer to this effect as the return to job tenure. For a worker i, $\beta_1 T_{ij(i,t)t}$ stands for the total wage gain from spending $T_{ij(i,t)t}$ years with firm $j(i,t)$. When the worker i is separated from the firm, he loses $\beta_1 T_{ij(i,t)t}$.

β_2 is the causal effect on wages of a worker spending one additional year in the labor market, holding the worker's job tenure constant, together with all the error components specified in equation (1.2). I refer to this effect as the return to the general labor market experiences.

Throughout the paper, the reported estimates of the return to tenure should be interpreted as the return to staying one more year with an employer, in excess of the return he gains from the additional year of experiences accumulated.

Identification assumption. I impose the following assumption to identify the key parameters of interest (β_1, β_2):

$$E[u_{it}|\mathbf{T}, \mathbf{M}, \mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\psi}] = 0 \tag{1.4}$$

where $\mathbf{T} \equiv \{T_{ij(i,t)t}\}_{i\in\{1,\ldots,N\},t\in\{1,\ldots,T\}}$, $\mathbf{M} \equiv \{M_{it}\}_{i\in\{1,\ldots,N\},t\in\{1,\ldots,T\}}$, $\mathbf{X} \equiv \{X_{it}\}_{i\in\{1,\ldots,N\},t\in\{1,\ldots,T\}}$, $\boldsymbol{\alpha} \equiv \{\alpha_i\}_{i\in\{1,\ldots,N\}}$, and $\boldsymbol{\psi} \equiv \{\psi_{jt}\}_{j\in\{1,\ldots,T\},t\in\{1,\ldots,T\}}$. This assumption imposes restrictions on workers' mobility which resembles the exogenous mobility assumption invoked in [1]. It entails two parts. First, the assumption excludes the situation where worker sorting and job mobility depend on worker-specific wage innovations that are not common within firms, u_{it}. Namely, the assumption states that workers do not move based on the realizations of u_{it}. This can be true in the case where u_{it} is paid to the worker independently of where the worker is employed. It is worth emphasizing that the assumption still allows flexible worker sorting patterns and job mobility based on the worker effects α_i, the firm effects $\psi_{j(i,t)t}$ and the worker's labor market experiences.[5] In particular, the assumption imposes no restriction on how long a worker chooses to work at a firm based on the worker and the firm effects. Importantly, workers are allowed to move across firms in response to firm-level wage shocks that are common to all workers within the origin (or destination) firm, ψ_{jt}. Second, under the assumption in equation (1.4), worker sorting and mobility are assumed to be orthogonal to worker-firm complementarities, $\phi_{ij(i,t)}$.[6] The worker-firm complementarity may be an important candidate for creating biases in return to tenure estimates, because workers tend to work longer at firms that turn out to be good matches, ceteris paribus.[7] Although not the focus of this study, it is important to keep in mind the two restrictions when interpreting the estimates of the return to tenure in this study.

1.4 Time-Invariant Heterogeneity

In this section, I first emphasize the importance of accounting for the worker and the firm permanent unobserved heterogeneity in estimating the return to tenure. By gradually allow-

5. Conditioning on the sequence of firm-year indicators $\boldsymbol{\psi}$ and the sequence of job tenure $\mathbf{T}$ is equivalent to conditioning on only the firm-year indicators $\boldsymbol{\psi}$.

6. If the worker-firm complementarities are additively separable in i and j, then the inclusion of worker fixed effects, α_i, and the time-invariant fixed effects, η_j, sufficiently controls for the interactive component. [11] provides identification of a framework that allows for worker-firm interactive effects and wage and mobility dynamics based on worker-level wage innovations u_{it}.

7. For example, see [4].

ing for richer unobserved heterogeneity from $\varepsilon_{ij(i,t)t}$ in equation (1.2), I discuss the intuitions, provide motivational evidence and show the estimation results of the biases. To help interpret the estimates across different model specifications in this section, I fix the estimation sample as the largest connected set of college male workers and firms.[8]

1.4.1 Worker Time-Invariant Heterogeneity

The goal in this subsection is to emphasize the importance of including worker unobserved heterogeneity in the estimation of the return to job tenure. I consider the estimation of the following special case of wage model (1.1):

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \gamma X_{it} + \delta_{r(i,t),t} + v_{it}$$

where it is assumed that $E[v_{it}|T_{ij(i,t)t}, M_{it}, X_{it}, \delta_{r(i,t),t}] = 0$. Importantly, it assumes that workers with high unobserved productivity are no more or less likely to have high tenure levels or labor market experiences. Moreover, any differences in firms' pay policies do not relate to the duration of job spells or workers' labor market experiences. Namely, α_i, $\eta_{j(i,t)}$ and $\tilde{\psi}_{j(i,t)t}$ are all assumed to be exogenous to a worker's job tenure and labor market experiences. I denote β_1 estimated from this model specification by β_1^{OLS}.

Intuition of the bias from the omission of α_i. One would expect a positive correlation between individual abilities of earning wages and job tenures. Less productive workers also tend to face a higher rate of layoffs and quits (perhaps due to workplace misconduct, and job mismatch). This possible correlation between worker unobserved productivity and job tenures create potential for an upward bias in β_1^{OLS}.

To show more rigorously how the correlation between α_i and job tenures create biases, I

8. I focus on a connected set of workers and firms because in the error model (1.2), both worker and firm fixed effects are present. Moreover, I make use of the estimated fixed effects to validate the intuition on the direction of biases. The concept of connected sets follows [1].

apply the Frisch–Waugh–Lovell (FWL) theorem to show the bias in β_1^{OLS}. The bias can be expressed as:[9]

$$\beta_1^{OLS} - \beta_1 = a_1^{OLS} + b_1^{OLS} + c_1^{OLS}$$

where a_1^{OLS}, b_1^{OLS} and c_1^{OLS} are the coefficients from the auxiliary regressions:

$$\begin{aligned} \alpha_i &= a_1^{OLS} T_{ij(i,t)t} + a_2^{OLS} M_{it} + \omega_{it}, \\ \eta_{j(i,t)} &= b_1^{OLS} T_{ij(i,t)t} + b_2^{OLS} M_{it} + \nu_{it} \\ \tilde{\psi}_{j(i,t)} &= c_1^{OLS} T_{ij(i,t)t} + c_2^{OLS} M_{it} + \iota_{it}. \end{aligned}$$

Compared with high-skilled workers who have the same labor market experiences, if low-skilled workers are more likely to be subject to job separations, perhaps because of the higher probability of workplace misconduct, then one should expect $a_1^{OLS} > 0$, and thus β_1^{OLS} is biased upward.

Suggestive evidence. Figure 1.1 shows suggestive evidence of the positive relation between worker unobserved productivity and job tenures. To construct this plot, I first rank workers based on their mean log real wages (net market×year effects and observed worker characteristics) and divide workers into 4 equally sized groups. Figure 1.1 displays the average job tenure levels of workers across these 4 quartile groups. A clear positive relationship emerges: those who earn more on average tend to be associated with higher job tenure levels in the sample. Figure 1.19 in the appendix repeats the same exercise but replaces the y-axis with job tenures residualized by general labor market experiences and worker characteristics. A similar positive relationship arises, suggesting that β_1^{OLS} is possibly biased upward due to the presence of worker permanent heterogeneity in productivity.

9. For simplicity, I abstract away the observed characteristics X_{it} and market × year fixed effects.

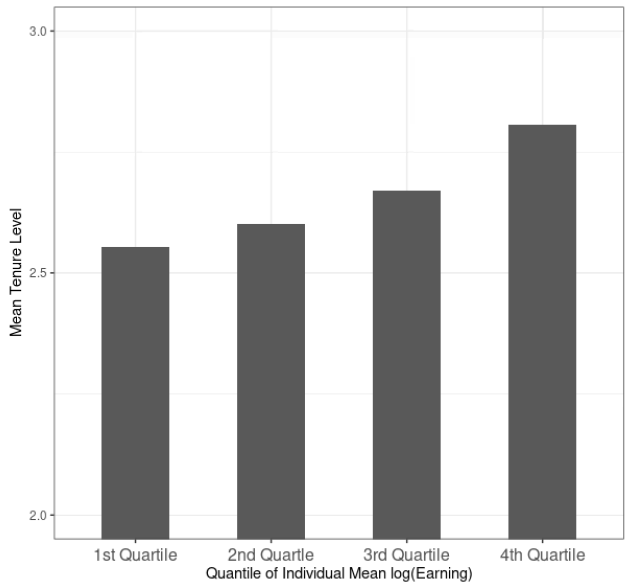


Figure 1.1: Covariance Between Individual Mean Wage and Tenure

Notes:
[1] The figure plots the average tenure level across 4 groups of workers ranked by their individual-specific mean wages. To construct this plot, I first residualize workers' wages by aggregate market ×year effects as well as worker characteristics. I then use the residualized wages to construct individual-specific mean wages for each worker in the sample. Next, I rank individuals based on their mean wages and partition individuals into 4 quartile groups.
[2] The result is based on a sample of college male workers in the private sector.

1.4.2 Bias from the Omission of Firm Heterogeneity

Motivated by the positive relationship between job tenures and a worker's mean wages, I now consider a more flexible wage model where worker unobserved heterogeneity is included but any unobserved variation in firm pay components is omitted:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \gamma X_{it} + \delta_{r(i,t),t} + \alpha_i + \tilde{v}_{it} \qquad \text{(FE-W)}$$

$$\tilde{v}_{it} = \eta_{j(i,t)} + \tilde{\psi}_{j(i,t)t} + u_{it}$$

where I assume that

$$E[\eta_{j(i,t)}|T_{ij(i,t)t}, M_{it}, X_{it}, \delta_{r(i,t),t}, \alpha_i] = 0$$
$$E[\tilde{\psi}_{j(i,t)t}|T_{ij(i,t)t}, M_{it}, X_{it}, \delta_{r(i,t),t}, \alpha_i] = 0.$$

I refer to β_1 that is estimated based on the above model specification as β_1^{FE-W}. The wage model assumes that, for a given worker, any differences in firms' pay policies do not induce differences in tenures. Importantly, this specification restricts that job moves are independent of the differences in firm fixed pay levels.

Intuition of the bias from the omission of η_j. The FE-W model is restrictive about the selection of job moves based on firm fixed pay levels. Holding workers' general labor market experiences and productivity constant, do higher-paying employment relationships (i.e., high η_j) tend to last longer than lower-paying ones (i.e., low η_j)? Theories provide competing insights on the sign of the covariance between job tenures and η_j. On the one hand, matching models and search theories (e.g., [12], [35], and [36]) suggest that when a worker is faced with a distribution of wages and when he is already being employed at a high-paying firm, the probability that he moves to some other job is low, because fewer potential outside offers can compete with his current job. Under this argument, high-paying firms tend to be associated with high tenure levels, creating an upward bias in the return to tenure estimates. I refer to this effect that positively correlates tenure levels and firms' average pay policies as the origin firm effect. On the other hand, [56] points out that workers voluntarily switch jobs for wage gains. Over careers, younger workers disproportionately sort from less productive to more productive firms ([30]). Consequently, later firms in a worker's career tend to be high-paying firms. Since they appear later in workers' careers, they may also be associated with lower observed tenure levels. Therefore, high-paying jobs can be correlated with low tenure levels. I call this effect the destination firm effect.

I use FWL theorem to demonstrate the two competing effects due to the presence of η_j. The bias in β_1^{FE-W} can be expressed as:

$$\beta_1^{FE-W} - \beta_1 = b_1^{FE-W} + c_1^{FE-W}$$

where b_1^{FE-W} and c_1^{FE-W} are the coefficients from the following auxiliary regressions:

$$\eta_{j(i,t)} = b_1^{FE-W} T_{ij(i,t)t} + b_2^{FE-W} M_{it} + b_3^{FE-W} \alpha_i + \omega_{it}$$
$$\tilde{\psi}_{j(i,t)t} = c_1^{FE-W} T_{ij(i,t)t} + c_2^{FE-W} M_{it} + c_3^{FE-W} \alpha_i + \iota it.$$

To help develop intuition, I use the first difference expression and write b_1^{FE-W} as:

$$b_1^{FE-W} = \frac{Cov(\mathbf{\Delta}\eta_{j(i,t)}, \mathbf{\Delta}T_{ij(i,t)t} - BLP(\mathbf{\Delta}T_{ij(i,t)t}|\mathbf{\Delta}M_{it}))}{Var(\mathbf{\Delta}T_{ij(i,t)t} - BLP(\mathbf{\Delta}T_{ij(i,t)t}|\mathbf{\Delta}M_{it}))}$$

where $\mathbf{\Delta}x_{it} \equiv x_{it} - x_{it-1}$, and $BLP(X|Y)$ stands for the best linear predictor of X by Y. To see what determines the sign of b_1^{FE-W}, it is useful to divide workers into two groups. At any time t, the employees at $j(i,t)$ can be partitioned into two mutually exclusive groups, (i) the incumbent workers who continued to work from $t-1$ to t with $j(i,t)$, i.e., job stayers with $j(i,t-1) = j(i,t)$, and (ii) the workers who are newly hired by $j(i,t)$ at t, i.e., job movers with $j(i,t-1) \neq j(i,t)$.

For job stayers, $\mathbf{\Delta}\eta_{j(i,t)}$ is zero, and $\mathbf{\Delta}T_{ij(i,t)t} = 1$ is deterministic. However, the between-firm differences in the firm pay premiums for a job mover, $\eta_{j(i,t)} - \eta_{j(i,t-1)}$, gives room for biases. With some simple algebra, the numerator of b_1^{FE-W} can be re-written as:

$$Cov(\boldsymbol{\Delta}\eta_{j(i,t)}, \boldsymbol{\Delta}T_{ij(i,t)t} - BLP(\boldsymbol{\Delta}T_{ij(i,t)t}|\boldsymbol{\Delta}M_{it})) = Pr(j(i,t-1) \neq j(i,t))\times$$

$$\{\underbrace{E[\eta_{j(i,t-1)}(T_{ij(i,t-1)t-1} - BLP(T_{ij(i,t-1)t-1}|\boldsymbol{\Delta}M_{it}))|j(i,t-1) \neq j(i,t)]}_{(a) \text{ origin firm effect}}$$

$$\underbrace{-E[\eta_{j(i,t)}(T_{ij(i,t-1)t-1} - BLP(T_{ij(i,t-1)t-1}|\boldsymbol{\Delta}M_{it}))|j(i,t-1) \neq j(i,t)]}_{(b) \text{ destination firm effect}}\}.$$

The expression indicates that the bias from the omission of firm fixed pay heterogeneity, b_1^{FE-W}, solely depends on job movers. It consists of three parts: (i) the gross probability of job moves, (ii) the relationship between the origin firm's pay level, $\eta_{j(i,t-1)}$, and a mover's foregone tenure from the origin firm, $T_{ij(i,t-1)t-1}$, and (iii) the relationship between the destination firm's pay level, $\eta_{j(i,t)}$, with the foregone tenure. Absent any job movers, the covariance shrinks to zero, but in the meantime, job tenure becomes perfectly colinear with the general labor market experiences. Holding the probability of job moves fixed, whether the FE-W estimator is unbiased depends on whether mobility is a function of origin and destination firms' pay levels and their relationships with the movers' foregone tenure.

Based on the above expression, two terms related to mobility determine the sign of the bias. Term (a) represents the origin firm effect. Suppose the origin firm $j(i,t-1)$ is high-paying. The mover from this high-paying firm is likely to have worked there for a long time because fewer outside options beat this employer's wage offer. Thus, a high η of the origin firm is likely to be related to high mover tenure levels, $T_{ij(i,t-1)t-1}$. Term (b) captures the destination firm effect. Workers typically progress to high-paying firms in their careers. To compensate for the wage loss from foregoing a high tenure level at the origin firm, one tends to ask for a high wage offer as compensation. As a result, movers' destination firms tend to be high-paying and associated with high foregone tenure levels. These two competing forces make the sign of b_1^{FE-W} ambiguous, depending on the relative magnitudes of the origin and the destination firm effects.

Suggestive evidence. Inspired by the above expression, I now show how movers' tenure levels vary in the data when they move across firms with different pay levels. In Figure 1.2, I compare the foregone tenure levels between two types of movers: upward movers and downward movers. To define upward (downward) movers, I first classify firms into two groups based on whether a firm's mean wage (residualized by worker characteristics including experiences, and market×year effects) exceeds the median of the firm pay distribution. A mover falls into the category of upward movers if he moves from a below-median firm group to an above-median firm group. The downward movers are defined if their movements are in the opposite direction. The previous expression of the bias predicts that the sign of b_1^{FE-W} is positive (i.e., the origin firm effect dominates) if movements that lead to larger wage gains (in terms of firm-specific pay levels) are associated with lower foregone tenures among movers. Figure 1.2 provides suggestive evidence for this prediction. Upward movers experience a slightly smaller drop in job tenures, around 0.2 years lower, than the downward movers.

In Figure 1.3, I show the differences in the destination and origin firms' mean wages by movers' foregone tenure levels. A similar but more clear negative relation arises between movers' foregone tenures and the wage gain in terms of firms' fixed pay levels. The patterns in Figure 1.2 and 1.3 both suggest that the origin firm effect is the dominating effect. Thus, one would expect an upward bias in the estimates obtained from the FE-W specification.

Estimation Results and Empirical Biases

Figure 1.4 presents the wage-tenure profiles estimated from three specifications: (i) the OLS specification, (ii) the FE-W specification where unobserved worker heterogeneity is controlled for, and (iii) the FE-TINV specification which accounts for both worker and firm

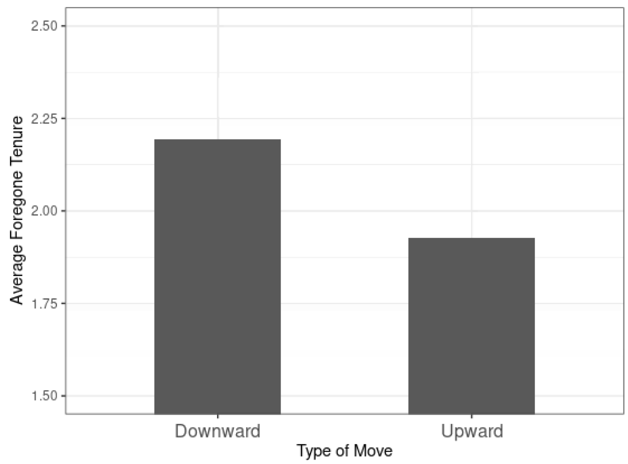


Figure 1.2: Change in Mean Foregone Tenure by Move Type: Upward versus Downward

Notes:
[1] The figure plots average movers' tenure at the origin firm across two types of job movers. Job movers are divided into two groups: upward movers who move from low-paying firms to high-paying firms, and downward movers who move from high-paying firms to low-paying firms. Firms are defined as high-paying if their firm-specific mean wages (residualized by worker characteristics and market × year effects) exceed the median of the distribution of firms' mean wages weighted by employment sizes. Similarly, firms are defined as low-paying if their firm-specific mean wages (residualized by worker characteristics and market × year effects) fall below the median of the wage distribution.
[2] The result is based on a sample of college male workers in the private sector.

time-invariant unobserved heterogeneity. Two results emerge.

First, across all three specifications, the wage growth from the return to firm tenure mainly comes from the first two years at a job, varying from 3.75 (FE-TINV) to 4.65 (OLS) log points per year. After two years of employment, the return to tenure flattens out and drops essentially to zero.[10] Note that the actual total labor market experiences are included in the wage model. Since the labor market experiences also contain the number of years a worker is employed at a firm, it is important to point out that the estimates of the return to tenure should be interpreted as the return to working one extra year at a firm in excess of the return to the general labor market experiences. Hence, the estimated wage-tenure profile implies that workers' wage growth comes mainly from the accumulation of general

10. The large wage growth from the initial years is unlikely to be driven solely by partial-year observations at the beginning of job spells because the wage measure in this study is hourly wage rather than total annual income.

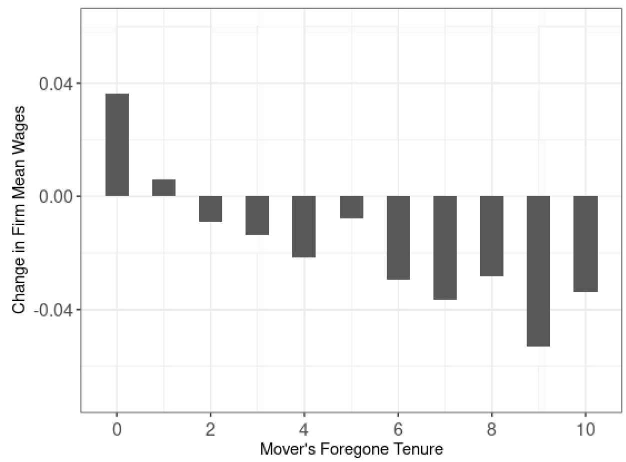


Figure 1.3: Differences in Destination and Origin Firm Mean Wages by Movers' Foregone Tenure

Notes:
[1] The y-axis is the average difference in the firm-specific mean wages between the destination and the origin firms for movers. The x-axis is the movers' tenure level at the origin firms.
[2] Firm-specific mean wages are constructed as the firm-specific average of residualized wages. Wages are residualized by workers' experiences, marital status, occupation, and market×year effects.
[3] The result is based on a sample of college males in the private sector.

labor market experiences two years after entry into a firm. Overall, the ten years of firm tenures lead to an increase in wages ranging from 5.5 (FE-TINV) to 8.7 (OLS) log points. These estimates are broadly in line with the findings in the literature.[11]

The second result pays attention to the comparative magnitudes of the return to tenure across the three specifications. The wage-tenure profile obtained from the FE-W specification lies uniformly below that from the OLS specification. The difference arises mainly from the first two years at the job, with a gap of around 0.74 log points each year. This gap between the OLS estimates and the FE-W estimates is consistent with the suggestive evidence of the positive covariance between tenures and individual productivity. The positive bias occurs likely because high productivity workers earn more and tend to hold longer jobs. As a result, workers with long tenure earn high wages in the cross-section.

11. In the reassessment of [56]'s and [3]'s approaches by [4], they estimate a return to 10-year firm tenures to be 10 log points with Topel's approach, and 4 log points with Altonji and Shakotko's approach, in the United States. [7] finds a much lower return, 3.6 log points, in the context of Norway. However, his specification of the wage model controls for firm permanent unobserved heterogeneity only.

The selection of job moves based on the differences in the firm pay levels further creates an upward bias in the return to job tenure estimates in FE-W. Once the selection along this dimension is taken into consideration, the return to tenure estimates are reduced by 0.16 log points per year in the first two years. This result implies that the origin firm effect on job moves indeed dominates the destination effect, in line with the suggestive evidence of the negative correlation of movers' between-firm wage gains and their foregone tenures. Compared to the bias created by unobserved worker heterogeneity, the magnitude of the correction by FE-TINV is much smaller.

In Figure 1.5, I plot the estimated return to experiences across the three specifications. Overall, the estimated return to ten years of experiences is broadly comparable with [56]'s and [3]'s estimates[12]. The estimated 10-year return ranges from 22.4 percent (FE-TINV) to 35.5 percent (OLS). Interestingly, the return to experiences drops significantly once firm-level heterogeneity in pay is controlled for. The fall in the FE-TINV estimates of the return to experience is likely a result of assortative worker sorting and workers' progression into better firms over careers.

Lastly, I report the empirical covariance of job tenures with the estimated fixed effects, $\hat{\alpha}_i$ and $\hat{\eta}_j$. The estimated covariance between $\hat{\alpha}_i$ and tenure is 0.1314, and the estimated $\hat{a}_1^{OLS} = 0.0021$. The empirical covariance between tenure and the estimated firm fixed effects $\hat{\eta}_j$ is 0.0275, with the estimated bias term $\hat{b}_1^{FE-W}$ to be 0.0010.

1.5 Time-Varying Firm Fixed Effects

1.5.1 Motivation

The inclusion of a firm-year component in the wage model is motivated by two pieces of evidence. First, a growing literature on the pass-through of firm-level shocks to workers'

12. [4] find an effect of 10-year experiences on wages to be 28.2 log points with Topel's approach, and 31.9 with Altonji and Shakotko's approach.

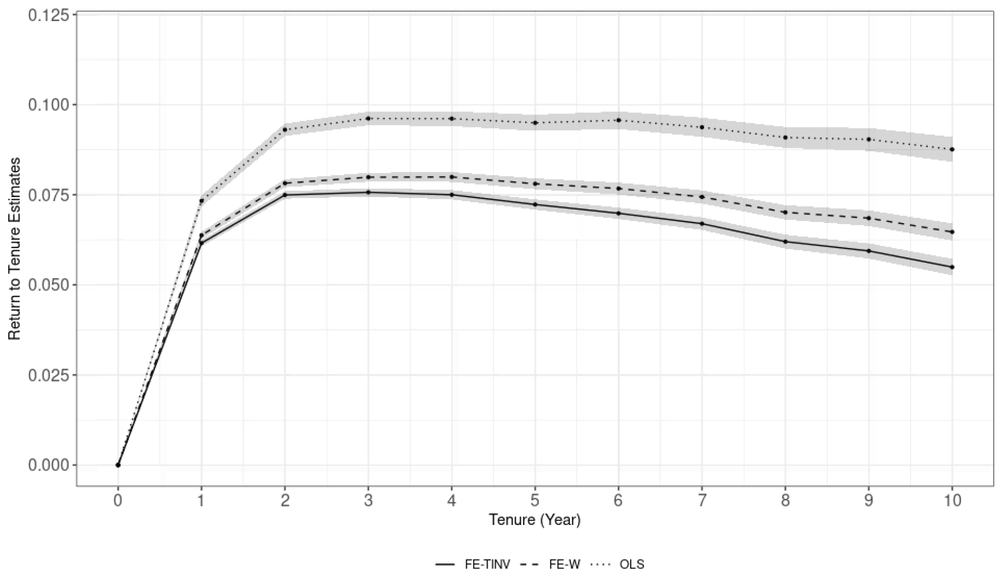


Figure 1.4: Return to Tenure Profiles Across Specifications

Notes:
[1] This plot shows the estimates of the return to tenure among college male workers.
[2] College male workers are defined as workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.1). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from the largest connected set of workers and firms for college male workers.
[5] The shaded area reports the 95% confidence bands.

wages invites attention to incorporating a time-varying firm pay component in modeling workers' wages. In an imperfectly competitive market, for example, due to the presence of search frictions, workers partially bear firm-level shocks from the product market since they cannot immediately secure outside job offers other than unemployment. As a result, within-firm changes in a worker's wages from one period to another (net the return to experiences) are a combination of the return to job tenure and a fraction of the firm-level shocks. For example, among firms that happen to receive a series of positive shocks, incumbent workers' wage growth at the firms is the sum of the true return to tenure and a fraction of the firms' positive shocks. Similarly, the observed changes in incumbent workers' wages can be less than the true return to job tenure at firms that happen to receive a series of negative shocks to their productivity.

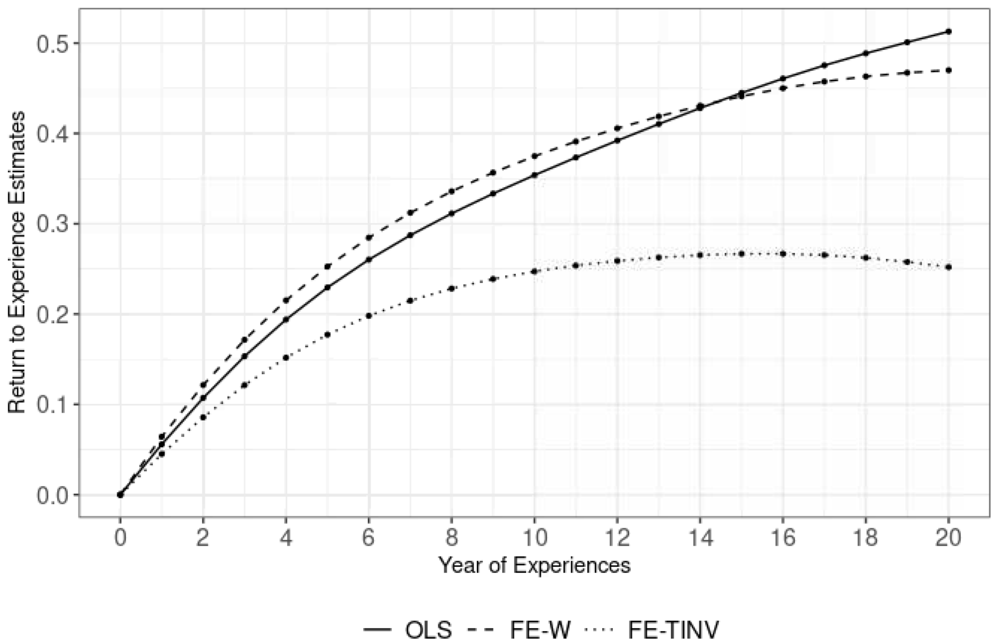


Figure 1.5: Return to Experiences Across Specifications

Notes:
[1] This plot shows the estimates of the return to experiences among college male workers.
[2] College male workers are defined as workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.1). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variables M_{it} enter the wage equation as 6th-degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from the largest connected set of workers and firms for college male workers.

The second motivation is that workers' job mobility decisions and job turnover can be a function of the aforementioned firm-level pay innovations. For instance, an incumbent worker may incorporate the anticipated growth in the firm pay premium from one period to another in his decision to continue to work with his current employer. Negative growth in the firm-specific pay can induce job separations because the lowered firm pay level makes outside offers more attractive. A more negative growth makes a worker more willing to give up the accumulated job tenures and seek outside options. Similarly, when a firm in an imperfectly competitive labor market receives better productivity shocks, it bids up wages to hire more labor, making itself a more attractive destination for job movers. The job movers who enter these firms reset their tenure level. Consequently, changes in a firm's productivity influence its employment composition, especially regarding employees' job tenure. This effect of the fluctuations in firms' pay policies on workers' mobility and job turnover creates a potential source for co-movement between the firm-level pay innovations and the workers' job tenure levels.

I provide empirical evidence to support that these forces are of empirical importance in the Norwegian context. First, I show that incumbent employees' wages co-move with their employers' productivity. Following the literature, I use log value-added to measure a firm's productivity (see [28], [43] and [27]). To demonstrate the worker-firm rent-sharing responses to firm-level shocks, I borrow the graphical illustration of the DID exercise from [43]. I examine whether there is co-movement between changes in the log value-added and changes in the incumbent workers' log wages between two groups of firms: those receiving above-median log value-added shocks, and those receiving below-median shocks. The exercise is carried out as follows. Denote firms' log value-added (net aggregate time effects) by y_{jt}. I construct the growth of log value-added from year $t-1$ to t for every firm j and calendar year t, and denote the growth by $\boldsymbol{\Delta} y_{jt}$. For every calendar year t (event year 0), I partition firms into high- and low-growth firms based on whether a firm's $\boldsymbol{\Delta} y_{jt}$ exceeds the median in the year-t distribution of $\boldsymbol{\Delta} y_{jt}$. Firms experiencing above-median growth in t fall into the treatment group in t while those whose growth falls below the median make up the control group.

Figure 1.6 presents the graphical evidence that incumbent workers' wages co-move with firm-level productivity shocks. The solid lines represent the difference between the treatment and control groups in the residualized log value-added 4 years before (event year < 0) and after (event year > 0) the event, weighted by firm employment sizes. The dotted lines trace the difference in the residualized log real wages of incumbent workers who are employed at the same firms 4 years before and after the event. In the event year 0, the treated firms experience a value-added growth around 49 percentage points larger than the control group. Correspondingly, workers in the treatment group experience an increase in wages by more than one percentage points than the control group. Overall, the evolution of the changes in the difference in incumbent workers' log wages mimic the changes in the firm's log value-added. The treated firms' log value-added stays high after the treatment, suggesting the presence of a permanent component in the log value-added process. The dips in event year -1 and 1 suggest the presence of mean reversion components. Similarly, the evolution in incumbent workers' wages points to the existence of a permanent and a mean-reverting component. However, wages seem to experience a lagged response to firms' log value-added growth.

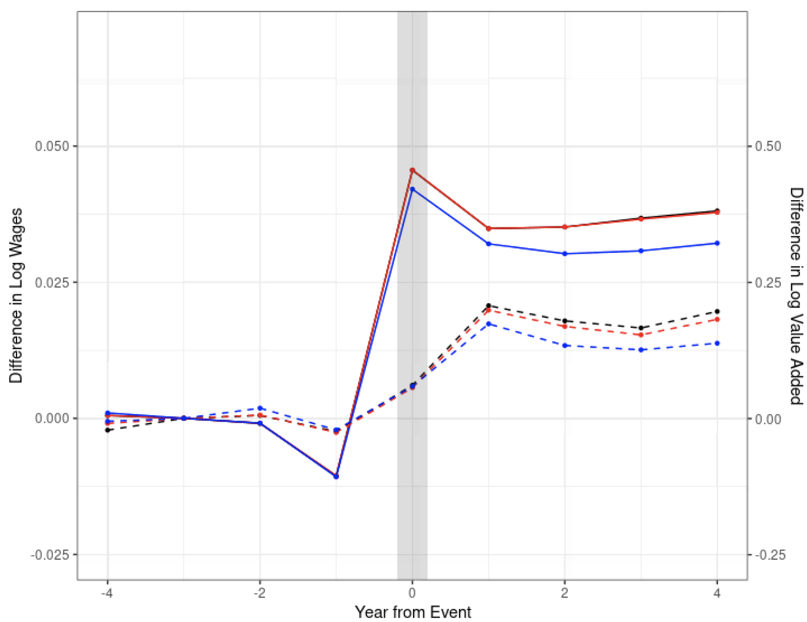


Figure 1.6: Co-movement Between Log Value-Added and Log Wage

Notes:
[1]The solid lines plot the differences in the evolution of firm's log value-added. The dotted line tracks the differences in the incumbent workers' log wages.
[2] The black lines represent log value-added (log wages) residualized by aggregate year effects. The red lines plot log value-added (log wages). The blue lines show the log value-added (log wages) residualized by market $\times$ year effects.
[3] Incumbent workers are required to stay at the same firm from event year -4 to event year 4.
[4] The differences in log value-added (log wages) between the treatment and control groups are normalized to 0 in event year -3.
[5] The sample of the college male workers is used for estimation.

Next, I present evidence that fluctuations in firms' productivity enter workers' job mobility along two dimensions: hiring and job separation (quitting/layoffs). I adapt an event study strategy similar to Figure 1.6. However, instead of examining the outcome of incumbent workers' wages, I focus on the employment share of newly hired workers and quitters. Figure 1.7 shows the result of the changes in the employment share of newly hired workers. The treatment group experiences a larger increase (around 1.8 log points) in the share of newly hired workers compared to the firms in the control group in event year 0. The increases in event year -1 and 1 suggest the presence of anticipation and persistent effects. The growth in firms' productivity drives up the hiring of workers and introduces an inflow of newly hired workers with the lowest tenure level into the firms. A negative relationship between job tenures and firm shocks can thus arise. Figure 1.8 examines the changes in the share of quitters/layoffs between the treatment and control groups as firms receive above-median log value-added shocks. The treated firms experience a lowered share of quitters/layoffs (0.012) compared to the control group, suggesting that workers have an increased propensity to continue working with their employers when the firms receive positive shocks. The dip in the share of quitters/layoffs prior to the year of the event suggests that workers possibly anticipate firm growth and incorporate that in their mobility decisions. While the results on the share of quitters are consistent with the prediction that workers incorporate current employers' future productivity growth in their mobility decision, it is also worthwhile to point out that the result also contains natural job separations (such as retirement) and involuntary job separations (such as layoffs). The former may cause less concern if the treatment and control groups experience similar trends in natural job separations.

As supporting evidence for the argument that workers anticipate future value-added shocks and incorporate it in their quitting decisions, I estimate the following linear probability model:

$$Quit_{ij(i,t)t} = \sum_{s=t-4}^{t+1} b_s \log(VA)_{j(i,t)s} + \delta X_{it} + v_{it}$$

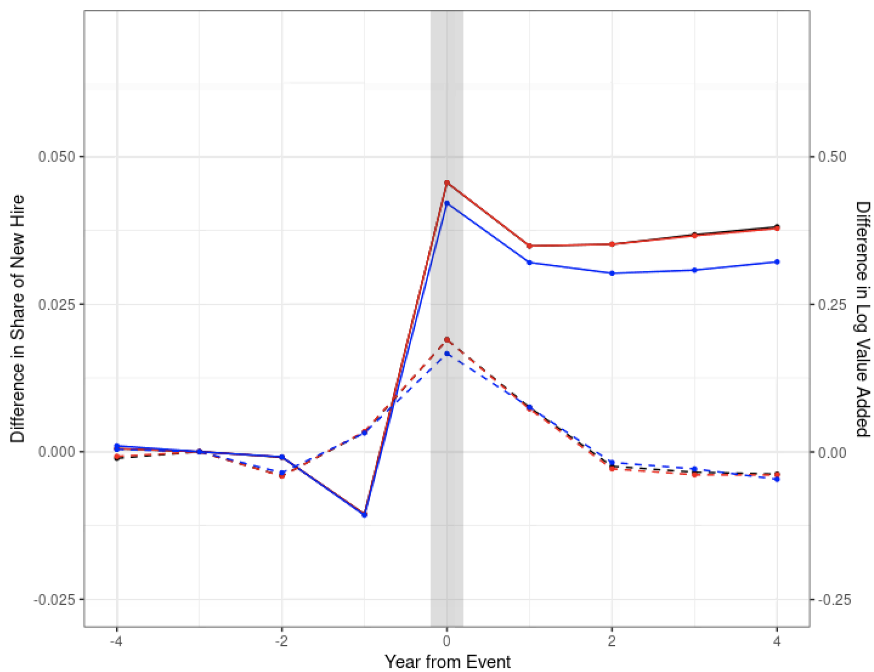


Figure 1.7: Differences in Log Value-Added and Share of New Hires

Notes:
[1] The solid lines plot the differences in the evolution of firm's log value-added. The dotted line tracks the differences in the employment share of newly hired workers.
[2] The black lines represent log value-added (share of new hires) residualized by aggregate year effects. The red lines plot log value-added (share of new hires). The blue lines show the log value-added (share of new hires) residualized by market × year effects.
[3] The differences in log value-added (share of new hires) between the treatment and control groups are normalized to 0 in event year -3.
[4] The sample of the college male workers is used for estimation.

where $Quit_{ij(i,t)t}$ is an indicator for whether worker i is separated from firm j at the end of t, and X_{it} contains worker characteristics, market×year, firm and worker fixed effects. Table 1.3 summarizes the results. Column (1) to (3) report the results where only the current and the past value-added are included in the estimation. Overall, a one percentage point decrease in the current value-added is associated with an increase of 0.0193 in the probability of job separation for workers. Column (4) and (5) additionally include the lead-year log value-added in the estimation. Holding everything else fixed, including the current and past log value-added, a one percentage point fall in the lead value-added is related to a 0.02 increase in the probability of job separation. The results offer supportive evidence that

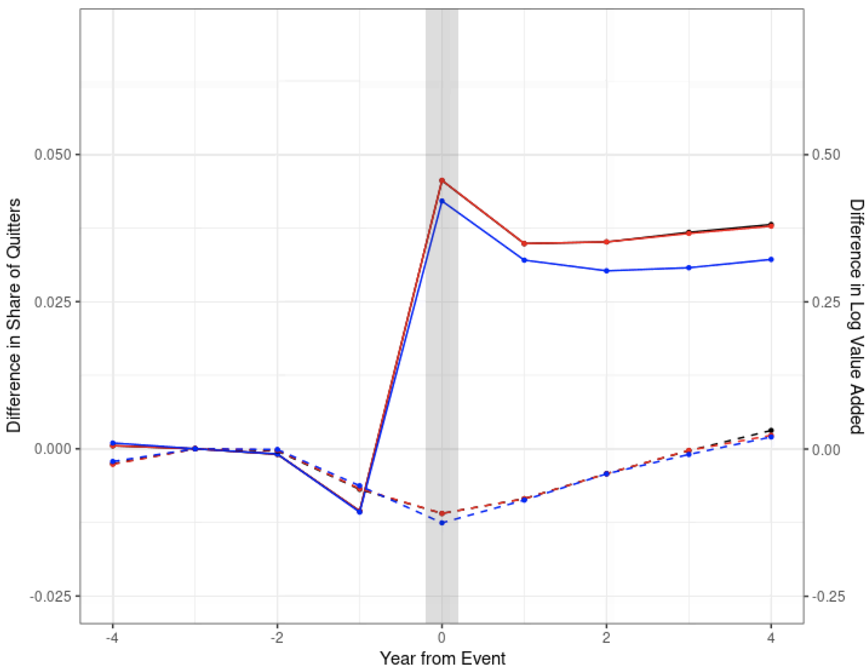


Figure 1.8: Differences in Log Value-Added and Share of Quitters

Notes:
[1] The solid lines plot the differences in the evolution of firms' log value-added. The dotted line tracks the differences in the employment share of quitters/layoffs.
[2] The black lines represent log value-added (share of quitters/layoffs) residualized by aggregate year effects. The red lines plot log value-added (share of quitters/layoffs). The blue lines show the log value-added (share of quitters/layoffs) residualized by market $\times$ year effects.
[3] The differences in log value-added (share of quitters/layoffs) between the treatment and control groups are normalized to 0 in event year -3.
[4] The sample of the college male workers is used for estimation.

workers anticipate future shocks to the current firm and incorporate them into the mobility decision.

Motivated by the changes in the employment composition due to hiring and quitting responses to firm productivity shocks, I examine how the average tenure levels at firms change with respect to these shocks. Figure 1.9 shows the differences in the average tenure levels between the treatment and control groups. Treated firms that receive high firm-level value-added shocks see an additionally lowered average tenure in the event year, and it continues to stay low in the subsequent years. Therefore, the net effect of the compositional changes in employment in response to value-added shock is negative on average tenures at

Vars. Quit	(1)	(2)	(3)	(4)	(5)
Current $\log(VA)_{jt}$	-0.0193	-0.0233	-0.0260	-0.0075	-0.0094
	(0.0004)	(0.0004)	(0.0005)	(0.0006)	(0.0006)
Lead $\log(VA)_{jt+1}$				$black{-0.0251}$	$black{-0.0227}$
				(0.0005)	(0.0006)
Lagged $\log(VA)_{j,t-1}$	0.0163	0.0064	0.0024	0.0076	0.0051
	(0.0004)	(0.0005)	(0.0006)	(0.0006)	(0.0006)
Lagged $\log(VA)_{j,t-2}$		0.0151	0.0039	0.0051	0.0027
		(0.0004)	(0.0006)	(0.0006)	(0.0006)
Lagged $\log(VA)_{j,t-3}$			0.0073	0.0100	0.0049
			(0.0006)	(0.0006)	(0.0006)
Lagged $\log(VA)_{j,t-4}$			0.0128	0.0147	0.0077
			(0.0005)	(0.0005)	(0.0006)
Firm FE	x	x	x	x	x
Worker FE					x
Market FE and worker X	x	x	x	x	x
Worker X	x	x	x	x	x
R^2	0.2593	0.2748	0.3024	0.3132	0.3749

Table 1.3: Quitting Based on Past, Current and Lead Value-Added

Notes:
[1] Worker X refers to worker-level controls including marital status, occupation, a 6th-degree polynomials in experiences, and worker fixed effects.
[2] The sample of the college male workers is used for estimation.

the firm level. This result suggests an inverse relationship between job tenures and firm-level shocks, calling for the examination of a time-varying firm component in the wage model.

1.5.2 Identification of Time-Varying Firm Effects and Estimation

Before presenting the estimation results of the return to job tenure and experiences, I discuss the identification of $\{\psi_{jt}\}_{j,t}$ and the estimation sample. $\{\psi_{jt}\}_{j,t}$ is a set of parameters of interest in this study, since they are key to understanding the determinants of the firm-level wage fluctuations and how they create biases in the previous estimators of the return to job tenure.

Identification of ψ_{jt}. The identification of ψ_{jt} is analogous to that of the AKM two-way fixed effects framework. The traditional AKM framework relies on the use of job movers be-

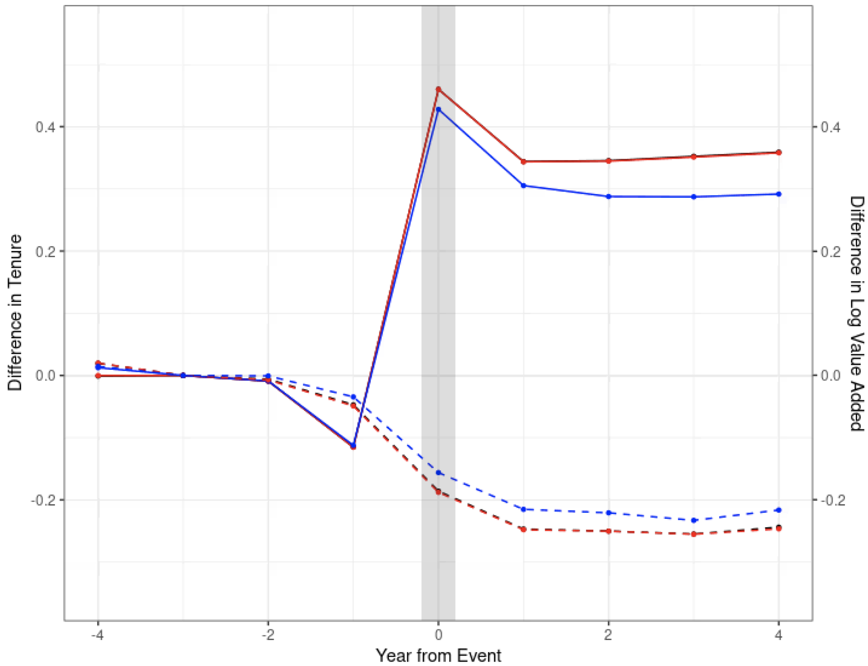


Figure 1.9: Differences in Average Tenure and Log Value-Added

Notes:
[1] The solid lines plot the differences in the evolution in firm's log value-added. The dotted line tracks the differences in firms' average tenure level.
[2] The black lines represent log value-added (firms' average tenure level) residualized by aggregate year effects. The red lines plot log value-added (firms' average tenure level). The blue lines show the log value-added (firms' average tenure level) residualized by market $\times$ year effects.
[3] The differences in log value-added (firms average tenure level) between the treatment and control groups are normalized to 0 in event year -3.
[4] The sample of the college male workers is used for estimation.

tween firms where firm identities are typically defined by physical establishments or plants. In particular, in the AKM framework, wages are additive in worker fixed effects, firm fixed effects, and an error term. The time-invariant firm effects can thus be identified as the differences in wages of job movers between physical establishments. Analogously, the identification of ψ_{jt} in the time-varying firm effect framework replies on "movers" between "firm$\times$year" pairs. Absent labor market experiences and job tenures, changes in the wages of job stayers identify the differences in the time-varying firm effects *within firms*. Differences in the wages of job movers across firms identify the difference in the firm effects *between firms* around the year of the job move. In this case, the worker and time-varying firm fixed effects can be simultaneously identified on a connected set of "movers" across firm×year pairs, up to one normalization. With the presence of both labor market experiences and job tenures, job stayers and job movers provide the variation that helps identify the worker and time-varying firm fixed effects.

For the remainder of this section and Section 1.6, I construct a connected set of firm×year pairs for the full estimation of model (1.1). I refer to the specification that controls for the time-varying firm fixed effects as FE-TV.

1.5.3 Results of Return to Job Tenure

Table 1.4 summarizes the estimation results across the 4 model specifications: OLS, FE-W, FE-TINV and FE-TV. The return to tenure profiles are plotted in Figure 1.10. Figure 1.11 shows the experience profiles. All estimates are obtained from the largest connected set (defined by firm×year pairs) of college male workers in the private sector in Norway.

Figure 1.10 reveals a key result in this study. Once the time-varying feature of firm-specific pay premiums is taken into account, a drastic increase in the return to job tenure arises, compared to the FE-TINV estimates. Relative to the FE-TINV specification, the FE-TV results push up the return to tenure by 0.49 log points per year in the first two years of employment. The 10-year return to tenure is estimated to be 7.44 log points. Moreover, the correction of the bias is so large that the return to tenure profiles now lies uniformly above the FE-W case. This downward bias in $\beta_1^{FE-TINV}$ suggests that job tenures and the fluctuations in the firm pay levels are strongly and inversely correlated.

In Figure 1.11, I plot the wage-experience profiles across the 4 specifications. Similar to the FE-TINV specification, the estimated return to experience is much lower than FE-W and OLS once firm-level heterogeneity is controlled for. On average, 10-years of experiences in the labor market lead to an increase in wages by around 25.4%. Based on the FE-TV results, although total labor market experiences are the main and more sustainable source of wage growth, the wage gain from the return to firm tenures in the first two years is comparable to the return to experiences.

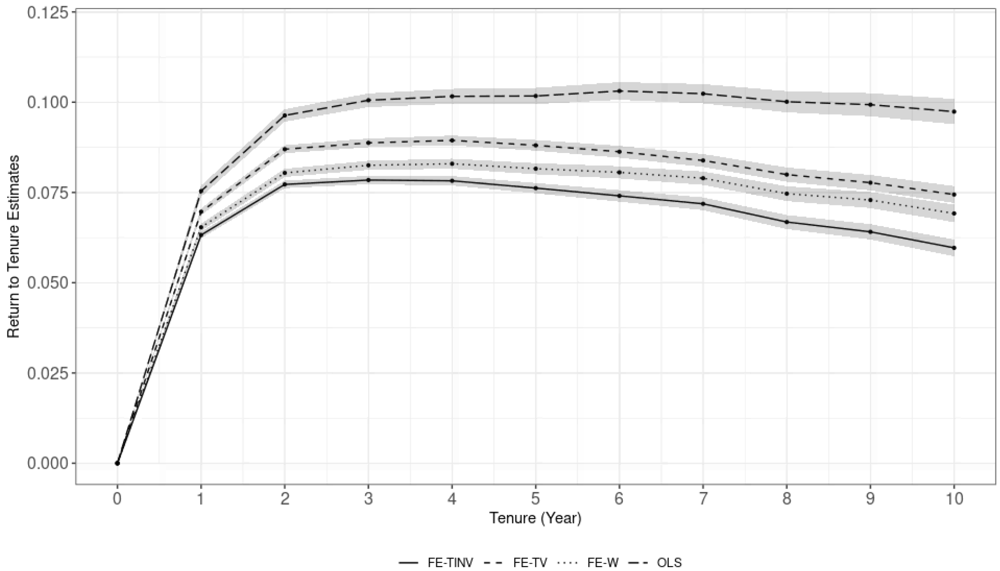


Figure 1.10: Return to Tenure Profile Among College Male

Notes:
[1] This plot shows the estimates of the return to tenure among college male workers.
[2] College male workers are defined as workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.1). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variables M_{it} enter the wage equation as 6th-degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from the largest connected set of workers and firm×years for college male workers.
[5] The shaded area reports the 95% confidence bands.

1.5.4 Instrumental Variable Strategies

I conclude this section by comparing the preferred estimates of the return to tenure to the literature. The literature on the return to firm tenure has provided various instrumental strategies to tackle the potential biases due to unobserved factors that affect mobility and wages simultaneously. Two influential and heatedly debated approaches stand out: [3]'s IV estimator and [56]'s two-stage IV estimator. Although the wage model considered in this paper is not entirely nested or nesting the wage model considered by these two approaches,

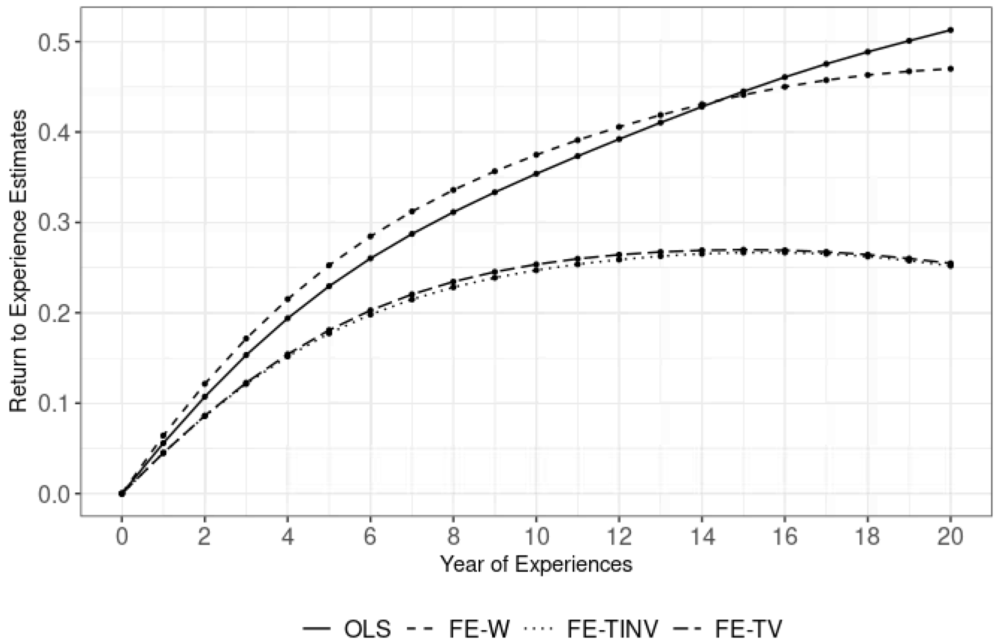


Figure 1.11: Return to Experiences Among College Male Workers

Notes:
[1] This plot shows the estimates of the return to experiences among college male workers.
[2] College male workers are defined as workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.1). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from the largest connected set of workers and firm×years for college male workers.

it is still helpful to compare their estimates to mine, and discuss the relative strength.

IV strategy I: Altonji and Shakotko's (1987) IV estimator. [3] considers the following wage model:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \alpha_i + \phi_{ij(i,t)} + \gamma X_{it} + \theta_t + u_{it} \tag{1.5}$$

where $\phi_{ij(i,t)}$ is the complementarity between a worker and a firm, θ_t is the aggregate year effects, and u_{it} is assumed to be independent of everything else in the model. They do not consider a firm-level time-varying pay component in workers' wages, $\tilde{\psi}_{jt}$. To address the

	OLS (1)	FE-W (2)	FE-TINV (3)	FE-TV (4)
2 Years of Tenure	0.0963	0.0804	0.0772	0.0870
	(0.0009)	(0.0006)	(0.0005)	(0.0006)
5 Years of Tenure	0.1017	0.0816	0.0762	0.0881
	(0.0011)	(0.0008)	(0.0007)	(0.0008)
10 Years of Tenure	0.0974	0.0692	0.0597	0.0744
	(0.0018)	(0.0012)	(0.0012)	(0.0012)
First 5-Year Average per Year (Experiences)	0.0461	0.0459	0.0321	0.0362
First 10-Year Average per Year (Experience)	0.0355	0.0341	0.0224	0.0254

Table 1.4: Returns to Tenure Among College Male Workers

Notes:
[1] This table reports the estimates of the returns to firm tenure and experiences among college male workers.
[2] College male workers are defined as workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.1). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from the largest connected set of workers and firm×years for college male workers.

biases in β_1 from $\phi_{ij(i,t)}$ and α_i, they use within-spell variation in $T_{ij(i,t)t}$ as an instrument for $T_{ij(i,t)t}$. Denote the within-spell variation by $DT_{ij(i,t)t} = T_{ij(i,t)t} - \bar{T}_{ij(i,t)}$. In the absence of $\tilde{\psi}_{j(i,t)t}$, the within-spell variation in tenure, $DT_{ij(i,t)t}$, is a valid instrument because it is orthogonal to the unobserved components that are fixed within job spells, in particular, α_i, and $\phi_{ij(i,t)}$. The Altonji and Shakotko's estimator is thus obtained by instrumenting $(T_{ij(i,t)t}, M_{it}, X_{it})$ by the vector of instruments $(DT_{ij(i,t)t}, M_{it}, X_{it})$. Denote the Altonji and Shakotko's estimators as β_1^{AS} and β_2^{AS}. Since the bias induced by the correlation between M_{it} and $(\phi_{ij(i,t)}, \alpha_i)$ still remains, the Altonji and Shakotko's estimators are still subject to potential biases, even without $\tilde{\psi}_{j(i,t)t}$.[13]

13. Under the argument that the general labor market experiences tend to be high for more productive workers, and for workers with better job matches, there can exist a positive bias in β_2^{AS}. And similarly, there may exist a negative bias in β_1^{AS}.

IV strategy 2: Topel (1991)'s two-Step first difference estimator. [56] considers a wage model similar to model (1.5). He breaks the estimation of β_1 and β_2 into 2 stages. In the first stage, he proposes to estimate the combined effect of the linear labor market experience β_2, the linear job tenure β_1 and the calendar time trend coefficient θ, using the implied wage growth model for job stayers s.t. $j(i,t-1) = j(i,t)$ (absent $\tilde{\psi}_{j(i,t)t}$) :

$$\mathbf{\Delta} w_{it} = \beta + \gamma \mathbf{\Delta} X_{it} + \mathbf{\Delta} u_{it}$$

where $\mathbf{\Delta} y_{it} \equiv y_{it} - y_{it-1}$. The parameter β is defined as $\beta = \beta_1 + \beta_2 + \theta$ where θ is the coefficient of the aggregate time trend. Topel notes that total labor market experiences can be re-expressed as the sum of the initial experience at the start of a job spell and the job tenure. Namely, $M_{it} = M0_{ij(i,t)t} + T_{ij(i,t)t}$ where $M0_{ij(i,t)t}$ is the labor market experiences worker i carries when entering firm $j(i,t)$. Hence, in order to separate β_1 from β_2 and θ, one can estimate the following second-stage wage model using OLS:

$$w_{it} - \hat{\beta} T_{ij(i,t)t} = \beta_2 M0_{ij(i,t)t} + \theta(t - T_{ij(i,t)t}) + \gamma X_{it} + e_{it} \tag{1.6}$$

where $\hat{\beta}$ is the first-stage estimator and $e_{it} = \epsilon_{ij(i,t)t} + t(\beta_0 - \hat{\beta}_0) + T_{ijt}(\beta - \hat{\beta})$. Thus, β_1 can be obtained as $\beta_1 = \hat{\beta} - \hat{\beta}_2 - \hat{\theta}$, once $(\hat{\theta}, \hat{\beta}_2)$ is obtained from equation (1.6). Although $\hat{\beta}$ is an unbiased estimator for the combined effect β, Topel's strategy is still subject to biases since the initial experience $M0_{ij(i,t)t}$ can be correlated with e_{it} where e_{it} contains α_i and $\phi_{ij(i,t)}$. As a result, the Topel estimators, β_1^T and β_2^T are not free from biases.

Results. I compare Topel's two-step first difference estimates, Altonji and Shakotko's IV estimates, and the FE-TV estimates on the sample of the college male workers in the private sector in Norway. Figure 1.12 presents the results of the return to tenure for the first ten years at a firm. Applied to the United State's data (PSID), Topel's and AS strategies give a 10-year tenure effect on log wages of 0.10 and 0.04, respectively ([4]). Consistently, in the Norwegian college male sample, the AS IV estimator produces a much smaller return to tenure than the estimates using Topel's approach. Yet, in the Norwegian context, both strategies yield estimates that are smaller than the estimates using the U.S data. The 10-year return to firm tenure is estimated to be 0.016 with AS approach, and 0.0590 using Topel's two-stage strategy. More interestingly, the Topel estimates are much closer to, though slightly smaller than, the FE-TV estimates on the Norwegian college male sample. When comparing the FE-TV estimates to AS and Topel estimates, one must be mindful that the FE-TV model is neither nested in nor nesting the model considered by Topel or AS. On the one hand, I assume away worker-firm complementarities in wages while their approach allows the presence of these complementarity effects (although their estimators are still subject to biases from the worker-firm complementarity through labor market experiences). On the other hand, their model does not allow a time-varying pay component at the firm level to enter a worker's job mobility. The results are summarized in Table 1.5.

1.6 Time-Varying Firm Effects and Source of Biases

In this section, I study how the time-varying firm effects create the downward bias. I start with deriving biases from the time-varying firm pay components and showing the ingredients that determine its sign. Next, I explore the determinants of this time-varying component in workers' wages. Lastly, I use various firm-level observables as controls to examine how much the inclusion of the firm-level observables helps reduce the bias.

1.6.1 Bias from Time-Varying Firm Fixed Effects

I now formally show the bias and the intuition from omitting the firm-level pay innovations in the FE-TINV estimates, $\beta_1^{FE-TINV}$.

Although the FE-TINV specification takes into consideration the selection of job moves

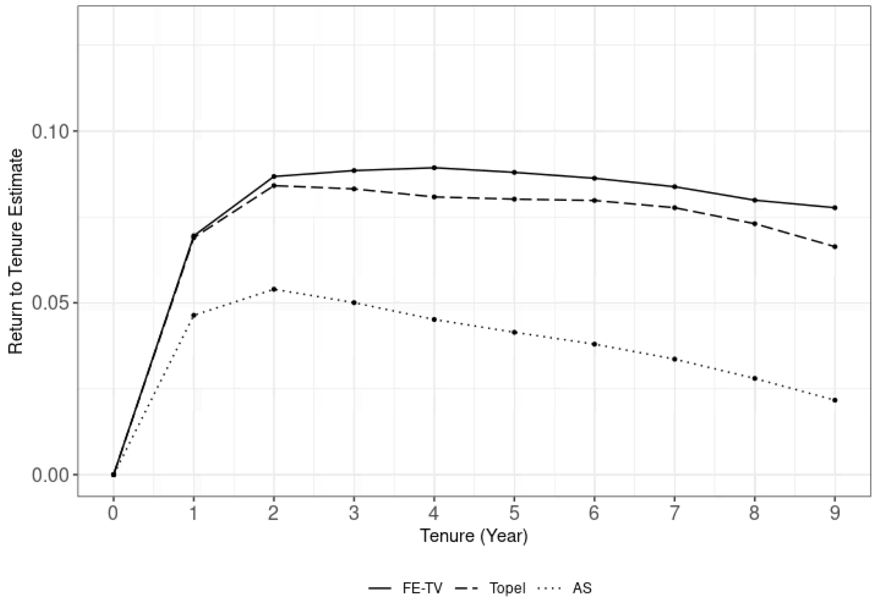


Figure 1.12: Comparison of IV Strategies and FE-TV Estimates

Notes:
[1] This figure shows the IV estimates using [56]'s and [3]'s IV estimators.
[2] "AS" refers to the estimates using [3]'s IV estimator. "Topel" refers to the estimates using [56]'s estimator. For both estimators, an 8th-degree polynomial in tenure, as well as a 6th-degree polynomial in experiences, are used.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation.

based on the origin and destination firms' fixed pay levels, it ignores when a worker moves regarding the draws of the origin and destination firms' time-varying pay innovations, i.e., when workers' tenure resets with respect to $\tilde{\psi}_{jt}$ and the foregone tenure $T_{ij(i,t-1)t-1}$. Before deriving the biases, I first confirm that this fluctuation in firms' pay levels over time indeed negatively correlates with workers' tenure in Table 1.6. Each column shows the results of regressing the estimated ψ_{jt} on a linear tenure term, with various controls. Job tenures and ψ_{jt} appear to be positively correlated in the cross-section (Column (1)), likely due to the fact that employment relationships on average last longer at higher-paying firms. However, once I zoom in on the within-firm variation in Column (4), the sign of their covariance flips to be negative. Within firms, tenure levels fall when ψ_{jt} increases, consistent with the graphical

	OLS (1)	FE-W (2)	FE-TINV (3)	FE-TV (4)	AS (5)	Topel (6)
2 Years of Tenure	0.0963	0.0804	0.0772	0.0870	0.0540	0.0841
	(0.0009)	(0.0006)	(0.0005)	(0.0006)		
5 Years of Tenure	0.1017	0.0816	0.0762	0.0881	0.0414	0.0802
	(0.0011)	(0.0008)	(0.0007)	(0.0008)		
10 Years of Tenure	0.0974	0.0692	0.0597	0.0744	0.0156	0.0590
	(0.0018)	(0.0012)	(0.0012)	(0.0012)		
5-Year Average per Year (Experiences)	0.0461	0.0459	0.0321	0.0362	0.0523	0.0374
	(0.0001)	(0.0001)	(0.0000)	(0.0000)		
10-Year Average per Year (Experiences)	0.0355	0.0341	0.0224	0.0254	0.0395	0.0320
	(0.0001)	(0.0001)	(0.0001)	(0.0000)		

Table 1.5: IV Estimates on College Male Workers

Notes:
[1] This table reports the IV estimates using [56]'s and [3]'s IV estimators.
[2] "AS" refers to the estimates using [3]'s IV estimator. "Topel" refers to the estimates using [56]'s estimator. For both estimators, an 8th-degree polynomial in tenure, as well as a 6th-degree polynomial in experiences, are used.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation, so the results are comparable across columns.

result in Figure 1.9.[14]

To formally show the bias, I consider the following simplified wage model where the observed characteristics X_{it} and the market×year fixed effects are abstracted away:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \alpha_i + \eta_{j(i,t)} + \tilde{\epsilon}_{ij(i,t)t} \qquad \text{(FE-TINV)}$$

$$\tilde{\epsilon}_{ij(i,t)t} = \tilde{\psi}_{j(i,t)t} + u_{it}.$$

The FE-TINV specification assumes that $E[\tilde{\epsilon}_{ij(i,t)t}|T_{ij(i,t)t}, M_{it}, \alpha_i, \eta_{j(i,t)}] = 0$. Applying the FWL theorem, the potential bias can be expressed as: $b_1^{FE-TINV} \equiv \beta_1^{FE-TINV} - \beta_1$,

14. Figure 1.20 in the appendix plots the estimated $\tilde{\psi}_{jt}$ across job tenures. A clear inverse relationship between job tenures and time-varying firm fixed effects confirms the negative bias in the return to tenure estimates.

Var. $\hat{\psi}_{jt}$	(1)	(2)	(3)	(4)
Tenure (linear)	0.0001	0.0001	-0.0012	-0.0011
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Year effects	x			
Market×year		x	x	x
Firm effects			x	x
Worker controls				x
Obs.	2,240,704	2,240,704	2,240,704	2,240,704

Table 1.6: Relationship Between ψ_{jt} and Tenure

Notes:
[1] This table reports the estimates of regressing the estimated $\hat{\psi}_{jt}$ on a linear tenure term.
[2] Worker controls include marital status, occupation dummies, a 6th degree polynomial in experiences and worker fixed effects.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation so that $\hat{\psi}_{jt}$ are observed.
[4] Firm employment sizes are used as weights in the estimation.

where $b_1^{FE-TINV}$ is the OLS coefficient in the following auxiliary regression:

$$\tilde{\psi}_{j(i,t)t} = b_1^{FE-TINV} T_{ij(i,t)t} + b_2^{FE-TINV} M_{it} + b_3^{FE-TINV} \alpha_i + b_4^{FE-TINV} \eta_{j(i,t)} + \nu_{it}. \tag{1.7}$$

To develop intuition for the sign of $b_1^{FE-TINV}$, I make use of the first difference estimator of b_1^{TINV} from equation (1.7):

$$\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t} = b_1^{FE-TINV} \boldsymbol{\Delta} T_{ij(i,t)t} + b_2^{FE-TINV} \boldsymbol{\Delta} M_{it} + b_4^{FE-TINV} \boldsymbol{\Delta}\eta_{j(i,t)} + \boldsymbol{\Delta}\nu_{it}$$

where $\boldsymbol{\Delta} x_{it} \equiv x_{it} - x_{it-1}$. Thus, $b_1^{FE-TINV}$ can be expressed as

$$b_1^{FE-TINV} = \frac{Cov(\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t}, \boldsymbol{\Delta} T_{ij(i,t)t} - BLP(\boldsymbol{\Delta} T_{ij(i,t)t} | \boldsymbol{\Delta} M_{it}, \boldsymbol{\Delta}\eta_{j(i,t)}))}{Var(\boldsymbol{\Delta} T_{ij(i,t)t} - BLP(\boldsymbol{\Delta} T_{ij(i,t)t} | \boldsymbol{\Delta} M_{it}, \boldsymbol{\Delta}\eta_{j(i,t)}))}. \tag{1.8}$$

To see what determines the sign of $b_1^{FE-TINV}$, it is again useful to divide workers at any time t into two groups: job stayers such that $j(i, t-1) = j(i, t)$, and job movers such that $j(i, t-1) \neq j(i, t)$.

For job stayers, $\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t}$ is the within-firm fluctuation in the firm pay premium, and $\boldsymbol{\Delta}T_{ij(i,t)t} = 1$ is deterministic. Therefore, job stayers' changes in tenure levels do not co-move with $\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t}$. However, the between-firm differences in the idiosyncratic firm pay premiums for a job mover, $\tilde{\psi}_{j(i,t)t} - \tilde{\psi}_{j(i,t-1)t-1}$, can be correlated with the mover's foregone job tenure, $T_{ij(i,t-1)t-1}$, for at least two reasons. The first reason relates to a worker's decision as to whether and when to continue the current employment relationship. Consider a worker who works at $j(i,t-1)$ at $t-1$. The idiosyncratic firm pay premium in the next period, $\tilde{\psi}_{j(i,t-1)t}$, may be (partially) observed by him, and therefore, the worker's decision to continue to work with $j(i,t-1)$ from $t-1$ to t is a function of the within-firm growth in the firm-specific pay innovations, $\tilde{\psi}_{j(i,t-1)t} - \tilde{\psi}_{j(i,t-1)t-1}$. The results in Table 1.3 lend support to this argument. An exceptionally low $\tilde{\psi}_{j(i,t-1)t} - \tilde{\psi}_{j(i,t-1)t-1}$ is likely to induce the worker to quit $j(i,t-1)$ early at $t-1$, everything else being equal. Moreover, this change needs to be low enough to justify the wage cost of quitting by a high-tenure worker. The second reason is about the draw of $\tilde{\psi}_{jt}$ at the mover's destination firm. A high counterfactual wage gain between the two firms, $\tilde{\psi}_{j(i,t)t} - \tilde{\psi}_{j(i,t-1)t}$, makes the job switch more attractive for a worker with a high foregone tenure level.

Formally, the numerator in $b_1^{FE-TINV}$ in equation (1.8) can be re-expressed as:

$$\begin{aligned}
&Cov(\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t}, \boldsymbol{\Delta}T_{ij(i,t)t} - BLP(\boldsymbol{\Delta}T_{ij(i,t)t}|\boldsymbol{\Delta}M_{it}, \boldsymbol{\Delta}\eta_{j(i,t)})) \\
&= -Pr(j(i,t-1) \neq j(i,t))\times \\
&\qquad \{\underbrace{E[(\tilde{\psi}_{j(i,t-1)t} - \tilde{\psi}_{j(i,t-1)t-1})\tilde{\tilde{T}}_{ij(i,t-1)t-1}|j(i,t-1) \neq j(i,t)]}_{(a)} + \\
&\qquad \underbrace{E[(\tilde{\psi}_{j(i,t)t} - \tilde{\psi}_{j(i,t-1)t})\tilde{\tilde{T}}_{ij(i,t-1)t-1}|j(i,t-1) \neq j(i,t)]}_{(b)}\}
\end{aligned}$$

where $\tilde{\tilde{T}}_{ij(i,t-1)t-1} \equiv T_{ij(i,t-1)t-1} - BLP(T_{ij(i,t-1)t-1}|\boldsymbol{\Delta}M_{it}, \boldsymbol{\Delta}\eta_{j(i,t)})$. The expression clearly shows that the bias solely depends on job movers, and there are two effects at play.

Term (a) is about how movers' foregone job tenures relate to the within-firm growth in $\tilde{\psi}_{jt}$ at the origin firm. Based on the previous arguments, a worker's decision to continue working with his employer may create a negative correlation between movers' foregone tenure levels and $\tilde{\psi}_{j(i,t-1)t} - \tilde{\psi}_{j(i,t-1)t-1}$, leading to an upward bias. Term (b) is about how foregone job tenures relate to the between-firm difference in the firm-level pay innovations for job movers. The previous arguments imply that the foregone tenures and the between-firm differences in $\tilde{\psi}_{jt}$ are likely to be positively correlated, which contributes to a downward bias in $\beta_1^{FE-TINV}$. Without knowing the relative magnitudes of the two effects, the sign of the bias is ambiguous.

I conduct the following decomposition exercise to support the above arguments. I compute the empirical counterparts of term (a) and (b) using the estimated $\tilde{\psi}_{jt}$'s from the FE-TV specification. The results are reported in Table 1.7. The signs of the empirical counterparts of term (a) and (b) conform with the predictions. The estimated between-firm effect dominates (around ten times larger than the within-firm growth effect). As a result, a negative bias arises in $\beta_1^{FE-TINV}$. The estimated $b_1^{FE-TINV}$ from the auxiliary regression (1.7) is $\hat{b}_1^{FE-TINV} = -0.0011$.

	Moments	Estimates
Total	$E[\boldsymbol{\Delta}\tilde{\psi}_{j(i,t)t}\tilde{\tilde{T}}_{ij(i,t-1)t-1} \vert j(i,t-1) \neq j(i,t)]$	0.0190
Between-Firm	$E[(\tilde{\psi}_{j(i,t)t} - \tilde{\psi}_{j(i,t-1)t})\tilde{\tilde{T}}_{ij(i,t-1)t-1} \vert j(i,t-1) \neq j(i,t)]$	0.0214
Within-Firm	$E[(\tilde{\psi}_{j(i,t-1)t} - \tilde{\psi}_{j(i,t-1)t-1})\tilde{\tilde{T}}_{ij(i,t-1)t-1} \vert j(i,t-1) \neq j(i,t)]$	-0.0024

Table 1.7: Moments of Tenure and $\tilde{\psi}_{jt}$ on Job Movers

Notes:
[1] $\tilde{\psi}$ are obtained from the estimation of model (1.1) on the largest connected set of workers and firm×years among college male workers. Movers are defined as the workers whose firm identity in year t differs from their firm identity in year $t+1$.
[2] $\tilde{\tilde{T}}_{ij(i,t-1),t-1}$ is obtained by regressing $T_{ij(i,t-1),t-1}$ on a constant, the estimated $\boldsymbol{\Delta}\eta_{j(i,t)}$ and $\boldsymbol{\Delta}X_{it}$.

1.6.2 Determinants of Time-Varying Firm Effects

Are the fluctuations in ψ_{jt} indeed associated with the fluctuations in the firm's performance in the product market? In this subsection, I address this question.

To answer how much ψ_{jt} can be explained by firm-level observables, especially measures of firm productivity, I explore the determinants of ψ_{jt} along two dimensions: the level of ψ_{jt} and the growth in ψ_{jt}. First, I project the estimated raw ψ_{jt} onto firm-level observables, using the following projection model:

$$\hat{\psi}_{jt} = Z_{jt}\delta^l + u_{jt}.$$

Z_{jt} is a vector of firm-level observables that are related to a firm's performance/productivity in the product market. Utilizing the rich firm balance sheets information available for this study, Z_{jt} includes firms' annual value-added, sales, operating revenues, firm employment sizes, and firm ages. I also include the market×year fixed effects in Z_{jt}.

In Table 1.8, I show the results of regressing the estimated ψ_{jt} on each of the firm-level variables and all together, weighted by the firm's employment size.[15] Table 1.12 in the appendix additionally includes firm fixed effects in the projection regression. The estimates in Table 1.8 and 1.12 reveal that the firm-level observables explain a large portion of the variation in ψ_{jt}. Even without the control for firm fixed effects, the R^2 of the projection regressions ranges from 0.211 to 0.368 in the univariate regressions. Due to the noises in the estimated ψ_{jt}'s, the R^2 should be interpreted as the lower bound of the explanatory power of these candidates. Among all the candidates, log employment sizes and log value-added have the highest explanation powers. With the inclusion of firm fixed effects, the explanation power jumps to as high as 0.864.

Firms' value-added and the estimated ψ_{jt} have an estimated elasticity of 0.0183, based

15. The projection results without using firm sizes as weights are included in the appendix.

on the regression results in Table 1.8. To help visualize how responsive ψ_{jt} is with respect to changes in a firm's value-added, I repeat the same event study exercise as in Section 1.5. Similar to the previous event study exercise, I partition firms into the treatment and control groups based on whether the growth in log value-added, $\boldsymbol{\Delta} y_{jt}$, exceeds the median of the year-specific distribution of $\boldsymbol{\Delta} y_{jt}$. I then plot the differences in the log value-added and the estimated ψ_{jt} between the treatment and control groups 4 years before and after the event. Figure 1.13 confirms that the evolution of ψ_{jt} mimics that of the log value-added, suggesting that ψ_{jt} partially tracks the firm's product market performances. Two additional interesting results are revealed. First, the fact that the dip in ψ_{jt} is much smaller in the event year -1 provides evidence that firms pass on persistent rather than transitory shocks to workers' wages. Moreover, the gradual increase in wages from event year 0 to event year 1 suggests a lagged effect of firm shocks on wages.

The previous projection results show that firms' productivity measures are good predictors for the cross-sectional variation in ψ_{jt} but not necessarily the within-firm changes in ψ_{jt}. To understand what drives the changes in ψ_{jt} within firms, I further consider the following projection model on growth in ψ_{jt}:

$$\boldsymbol{\Delta} \hat{\psi}_{jt} = \boldsymbol{\Delta} Z_{jt} \delta^{d} + u_{jt}$$

where Z_{jt} includes the same set of variables except for firm ages. Table 1.9 presents the results of the projection model. Among all the candidates, value-added and firm employment sizes continue to have the highest explanatory powers for $\boldsymbol{\Delta} \psi_{jt}$, with qualitatively similar magnitudes of elasticities. However, the R^2 shrinks to around 0.05. While firm-level productivity measures such as log value-added are good predictors for the level of ψ_{jt}, they capture mostly the cross-sectional variations and thus have a lower explanatory power, though non-trivial, in predicting the dynamics of this spatial component in wages. Put it another way, while the firm-specific pay components ψ_{jt} indeed partially reflect firms' productivity dif-

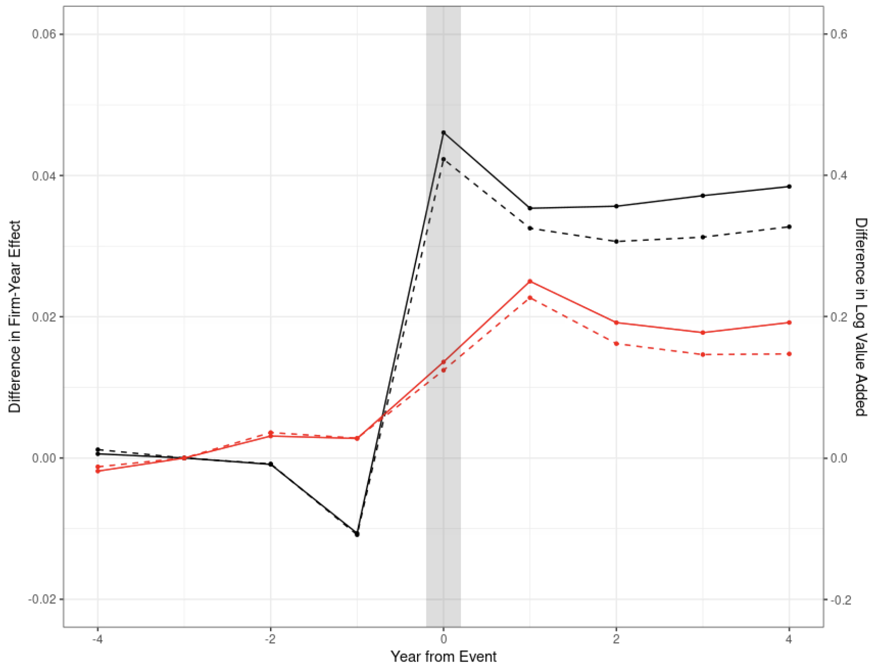


Figure 1.13: Difference in Log Value-Added and ψ_{jt}

Notes:
[1] In this plot, the black lines trace the differences in the evolution of firms' log value-added. The red line tracks the estimated ψ_{jt}. The solid lines represent raw log value-added and raw ψ_{jt}. The dotted lines represent the variables net market×year effects. The differences in the outcome between the treatment and control groups are normalized to 0 in event year -3.
[2] Incumbent workers are required to stay at the same firm from event year -4 to the event year 4.
[3] The differences in log value-added (ψ_{jt}) between the treatment and control groups are normalized to 0 in event year -3.
[4] The largest connected sample of workers and firm×years among college male workers is used for estimation.

ferences and fluctuations, there is still a large portion in ψ_{jt}'s that remained unexplained. Several factors can explain the lowered R^2. First, the noises in the estimated ψ_{jt} are large due to the large number of parameters estimated on a limited set of observations.[16] Second, the measures of log value-added (and the other firm-level variables) can be subject to large noises too. Moreover, firms pass permanent shocks to workers' wages, rather than idiosyncratic shocks ([43]). Within-firm variations in log value-added (and the other productivity measures) likely contain a large portion of idiosyncratic shocks, and such variations are not reflected in workers' wage variations ψ_{jt}. Last but not least, the firm time-varying pay com-

16. It is well documented that the fixed effects from the AKM framework are subject to large noises due to the large number of parameters estimated on the limited number of movers. See for [10] for details.

ponents ψ_{jt} is also likely to contain other firm-level shocks, such as changes in amenities, which affect firm employment compositions and wages simultaneously.

Var. Raw $\hat{\psi}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(VA)	0.0243							0.0292	0.0203
	(0.0001)							(0.0009)	(0.0020)
Lagged log(VA)		0.0218							0.0034
		(0.0001)							(0.0020)
Log(sales)			0.0226					0.0138	0.0222
			(0.0001)					(0.0009)	(0.0023)
Log(revenue)				0.0200				0.0056	0.0001
				(0.0001)				(0.0004)	(0.0005)
Firm age					0.0064			0.0001	0.0005
					(0.0001)			(0.0001)	(0.0001)
Log(firm size)						0.0250		0.0026	0.0342
						(0.0001)		(0.0005)	(0.0020)
Lagged log(firm size)							0.0224		-0.0325
							(0.0001)		(0.0020)
R^2	0.356	0.359	0.355	0.368	0.211	0.344	0.353	0.289	0.308
Obs.	1,745,011	1,645,492	1,902,336	1,578,383	717283	2,240,704	2,195,416	441,967	331,338
Mkt×year	x	x	x	x	x	x	x	x	x
Weights	x	x	x	x	x	x	x	x	x

Table 1.8: Firm Observables and ψ_{jt} on Unbalanced Sample, Weighted by Employment Size
Notes:
[1] This table reports the estimates of regressing the estimated $\hat{\psi}_{jt}$ on the firm-level variables.
[2] The variable "VA" refers to value-added, "revenue" refers to operating revenues.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation so that $\hat{\psi}_{jt}$ are observed. Since value-added, sales and operating revenues are not observed or are negative for some observations, the estimation samples vary across Column (1) to (9).
[4] Firm employment sizes are used as weights in the estimation.

1.6.3 Firm-Level Observables as Exogenous Controls

Having established that ψ_{jt} is correlated with the firm-level characteristics that measure firm performances, I investigate how much bias in $\beta_1^{FE-TINV}$ can be corrected by the inclusion of these firm observables.

I first gradually add firm-level observables to the regressions of $\hat{\psi}_{jt}$ on a linear tenure term:

$$\hat{\psi}_{j(i,t)t} = \delta^T T_{ij(i,t)t} + \delta^Z Z_{j(i,t)t} + \gamma X_{it} + \alpha_i + \eta_{j(i,t)} + \nu_{it}$$

Var. Raw $\boldsymbol{\Delta}\hat{\psi}$	(1)	(2)	(3)	(4)	(5)
Δlog(value-added)	0.0344				0.0205
	(0.0007)				(0.0014)
Δlog(sales)		0.0285			0.0285
		(0.0006)			(0.0018)
Δlog(revenue)			0.0034		-0.0039
			(0.0003)		(0.0004)
Δlog(firm size)				0.0402	0.0194
				(0.0012)	(0.0018)
R^2	0.062	0.055	0.060	0.046	0.073
N. observations	285,334	312,976	224,769	388,753	200,567
Mkt×year	x	x	x	x	x
Weights	x	x	x	x	x

Table 1.9: Firm Observables and $\Delta\psi_{jt}$ on Unbalanced Sample, Weighted by Employment Size

Notes:
[1] This table reports the estimates of regressing the estimated $\Delta\hat{\psi}_{jt}$ on the firm-level variables. $\Delta\hat{\psi}_{jt}$ is defined as $\hat{\psi}_{jt} - \hat{\psi}_{jt-1}$.
[2] The variable "VA" refers to value-added, and "revenue" refers to operating revenues.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation. Since value-added, sales, and operating revenues are not observed or are negative for some observations, the estimation samples vary across Column (1) to (5).
[4] Firm employment sizes are used as weights in the estimation.

where $Z_{j(i,t)t}$ includes value-added, sales, operating revenues, firm sizes, as well as the employment share of newly hired workers, and the employment share of quitters. The last two variables are additionally included, motivated by the decomposition exercise of the bias and the result that the hiring and quitting behaviors in relation to firm time-varying pay innovations play an important role in generating the bias. The share of newly hired workers is defined as the total number of newly hired workers at firm j in year t divided by the total employment of firm j in year t. Similarly, the share of quitters is defined as the total number of workers who leave firm j in year t by j's total employment in t. Lastly, X_{it} is a vector of worker controls, and including labor market experiences.

Table 1.13 reports the regression results. The coefficients of the linear tenure term increase back to almost 0 as the variation in ψ_{jt} is gradually absorbed by richer sets of firm-level observables, leaving limited room for correlation between job tenures and ψ_{jt}. It is noteworthy that the coefficients of job tenures increase the most with the inclusion of two types of firm-level variables: log value-added, and the share of newly hired workers. This is consistent with the findings that ψ_{jt} co-moves with firm performance via the worker-firm rent-sharing mechanism, and that between-firm growth in ψ_{jt} plays a dominating role in driving the resetting process of tenures among movers.

Motivated by the results in Table 1.13, I include a set of firm characteristics as exogenous firm-level controls in the following wage model:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \gamma X_{it} + \kappa Z_{j(i,t)t} + \alpha_i + \eta_{j(i,t)} + \delta_{r(i,t)t} + u_{it} \tag{1.9}$$

where $Z_{j(i,t)t}$ is a vector of firm-level measures, including log value-added, log firm employment sizes, and employment shares of new hires and quitters. To more flexibly capture lagged responses and growth, I include current and lagged variables (up to $t-2$) for each type of firm-level measure considered above. The objective is to examine how much these firm-level observables help correct the negative bias in the wage-tenure profile, relative to the FE-TINV estimates.

The results are shown in Figure 1.14 and summarized in Table 1.14. The black and red lines trace the FE-TV and the FE-TINV return to tenure estimates, respectively. The blue line represents the estimates of model (1.9) with various choices of Z_{jt}'s. Interestingly, neither log value-added nor log employment sizes alone help eliminate much of the bias stemming from the selection of job moves based on firm-level pay innovations over time. Instead, the firm employment shares of newly hired workers perform quite well in eliminating the downward bias. This is consistent with the previous finding that firm productivity measures such as log value-added explain some but not greatly of the within-firm dynamics of ψ_{jt}. As a result, log value-added (or log firm size) itself does not help much reduce the bias. However, the variation in the firm employment share of newly hired workers seems to be a good proxy for the firm-level pay innovations and the relation with workers' tenure levels. In this sense, the firm employment share of newly hired workers may be viewed as reflecting booms and busts at the firm level, and capturing the co-movement between tenure and firm-level shocks that attract workers and drive up wages simultaneously (e.g., firm amenities).

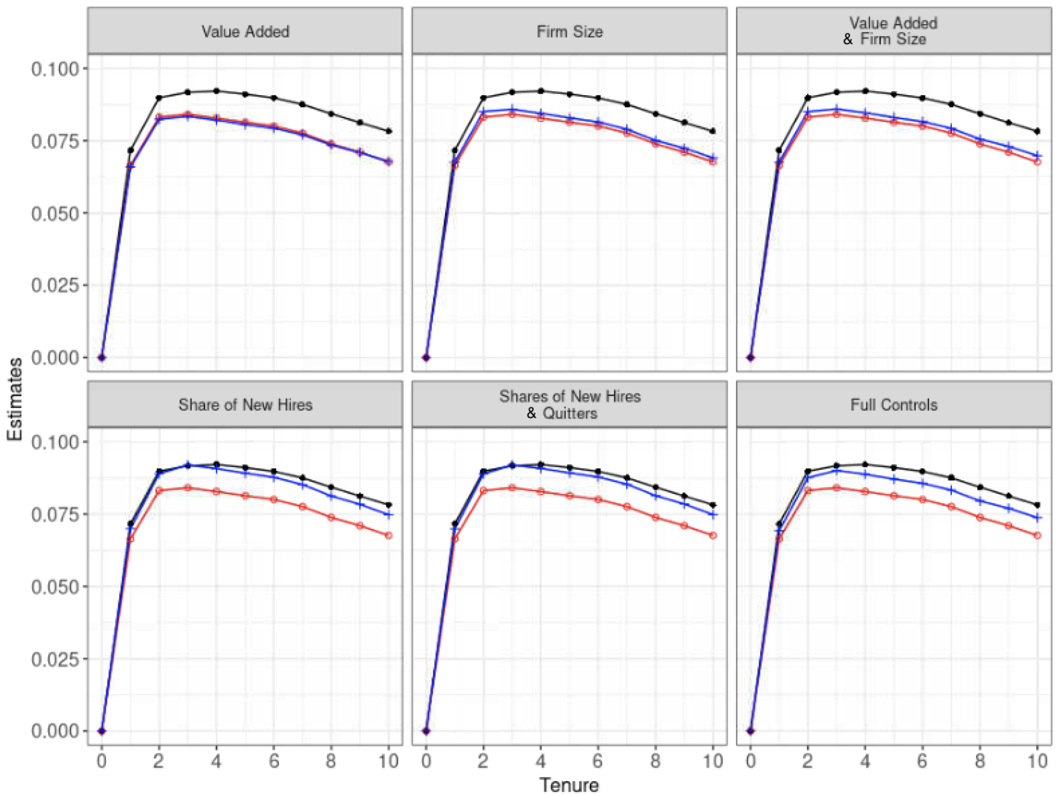


Figure 1.14: Correction of Bias with Firm-Level Observables

Notes:
[1] This plot compares the estimates of the return to job tenure using the wage model (1.9). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variable M_{it} enters the wage equation as a 6th degree polynomial. X_{it} consists of marital status, and occupation dummies.
[2] $Z_{j(i,t)}$ is a set of firm-level variables. For Panel "Value-Added", Z_{jt} includes the log value-added in period t, $t-1$ and t. For Panel "Firm Size", Z_{jt} includes the log firm size in period t, $t-1$ and t. For Panel "Value-Added and Firm Size", Z_{jt} includes the log value-added and log firm size in period t, $t-1$ and t. For Panel "Share of New Hires", Z_{jt} includes the number of newly hired workers divided by the employment size in period t, $t-1$ and t. For Panel "Share of New Hires and Quitters", Z_{jt} includes the share of new hires and the number of quitters/layoffs divided by the employment size ain period t, $t-1$ and t. For Panel "Full", Z_{jt} includes the log value-added, log firm size, the share of new hires and the share of quitters in period t, $t-1$ and t.
[3] The red line represents the FE-TINV estimates. The black line represents the FE-TV estimates. The blue line represents the estimates from model (1.9) with Z_{jt} corresponding to each panel.
[4] All estimates are obtained from the largest connected set of worker and firm×year where all variables in Z_{jt} are observed (and log value-added is greater than 0).

1.7 Monte Carlo Simulations

In this section, I conduct Monte Carlo simulation exercises to verify whether and in which direction the OLS, FE-W, and FE-TINV specifications produce biased estimates under different mobility assumptions. I first specify the data generating process. Importantly, I describe in detail the various mobility models considered for the Monte Carlo exercise. Next, I evaluate the directions of biases under the 4 specifications on the simulated data.

1.7.1 Simulation Setup

In the simulation exercise, I generate a network of N workers, J firms over T periods. I assume that there are n_l types of workers and n_k classes of firms. Each worker is given a draw of worker type l (and consequently a realization of worker-specific productivity value α_l). Each firm is given a firm type k (and thus a realization of firm-specific pay level, η_k, and a vector of time-varying pay components $(\tilde{\psi}_{k1}, ..., \tilde{\psi}_{kT})$).

For a worker i working at some firm j with $(\alpha_i, \eta_j, \tilde{\psi}_{jt}, T_{ijt}, u_{it})$, the log wage is determined by

$$w(\alpha_i, \eta_j, \tilde{\psi}_{jt}, T_{ijt}, u_{it}) = \beta_1 T_{ijt} + \alpha_i + \eta_j + \tilde{\psi}_{jt} + u_{it}.$$

Note that $w(\alpha_i, \eta_j, \tilde{\psi}_{jt}, T_{ijt}, u_{it})$ is the potential wage.

Now, I describe how job movement is generated. As discussed, job mobility is crucial in determining the bias in the return to job tenure. In particular, worker productivity, origin and destination firms' fixed pay levels, and within-firm and between-firm differences in $\tilde{\psi}_{jt}$, together with a mover's foregone tenure, jointly enter the mobility function and govern the sign of the bias. I consider four mobility models, from the most restrictive to the least restrictive.

Case (1): Random mobility. At the beginning of the first period $t = 1$, each worker i draws his worker type l, firm $j(i, 1)$, and an initial tenure level $T_{i,1}$, which are mutually independent of each other. Independently of everything else, he draws the worker innovation $u_{i,1}$ from $F_u(u)$.

At the beginning of each period t, the worker draws a potential firm $j^*(i, t)$ from $\{1, ..., J\}$. He also independently draws worker innovation u_{it} from $F_u(u)$. Then, he moves to the potential firm and resets the tenure level with an exogenously set probability p. In this case, job mobility is completely random. In particular, it is independent of a mover's foregone tenure level at the origin firm. One would expect that the return to tenure estimator using OLS is unbiased.

Case (2): Mobility based on worker type. At the beginning of the first period $t = 1$, each worker i draws his worker type l, firm $j(i, 1)$, and an initial tenure level $T_{i,1}$, which are mutually independent of each other. Independently of everything else, he draws the worker innovation $u_{i,1}$ from $F_u(u)$.

At the beginning of each period t, the worker of type l has probability $p(\alpha_l)$ to draw a potential firm $j^*(i, t)$ from $\{1, ..., J\}$. He also draws worker innovation u_{it} from $F_u(u)$. Then, he moves to the potential firm and resets the tenure level. I assume $p_2(\alpha_l)$ is decreasing in α's. In this way, a worker with lower α is faced with a higher probability of job separations and thus lower average tenure. But conditional on worker type, job mobility is random. One would expect that the return to tenure estimator using OLS is biased upward but the FE-W estimator is unbiased.

Case (3): Mobility based on firm time-invariant effects. At the beginning of the first period $t = 1$, each worker i draws his worker type l, firm $j(i, 1)$, and an initial tenure level $T_{i,1}$ from a joint distribution F. Independently of everything else, he draws the worker innovation $u_{i,1}$ from $F_u(u)$.

At the beginning of each period t, th worker of type l has probability $p_3(\alpha_l)$ to draw a potential firm $j^*(i,t)$ from $\{1,...,J\}$. He draws worker innovation u_{it} from $F_u(u)$. Then, he sees the firm-specific wage determinants from both the current firm, $(\eta_{j(i,t-1)}, T_{ij(i,t-1)t-1})$, and his potential firm $\eta_{j^*(i,t)}$. Note that the worker only sees the time-invariant pay levels η_j, not the time-varying components $\tilde{\psi}_{jt}$. The return to spending one extra year at the current firm, β_1, is known to him. He then makes the mobility decision based on the comparison of the two wage offers. In particular, conditional on receiving potential job offers, I assume mobility function $m(\alpha_i, \eta_{j(i,t-1)}, \tilde{\psi}_{j(i,t-1),t}, \eta_{j^*(i,t)}, \tilde{\psi}_{j^*(i,t),t}, T_{ij(i,t-1)t-1}, u_{it}|\beta_1) \equiv m(\eta_{j(i,t-1)}, \eta_{j^*(i,t)}, T_{ij(i,t-1)t-1}|\beta_1)$ as follows:

$$\begin{aligned} j(i,t) = j^*(i,t) \qquad & \Longleftrightarrow \underbrace{\beta_1(T_{ij(i,t-1)t-1} + 1) + \alpha_i + \eta_{j(i,t-1)} + u_{it}}_{\text{current firm}} < \\ & \qquad\qquad \underbrace{\alpha_i + \eta_{j^*(i,t)} + u_{it}}_{\text{potential firm}} + \nu_{it} \\ & \Longleftrightarrow \beta_1(T_{ij(i,t-1)t-1} + 1) + \eta_{j(i,t-1)} < \eta_{j^*(i,t)} + \nu_{it} \end{aligned}$$

where $\nu_{it} \sim \mathcal{N}(0, \sigma_\nu^2)$ is a normally distributed noise independent of everything else. Note that the worker-specific components (α_i, u_{it}) are dropped because they are paid regardless of where the worker is employed. Note also that, modeled this way, the worker's mobility is based on firms' fixed pay levels at the origin and the destination firms; however, he does not observe (or does not take into consideration) the time-varying firm pay component from either firm.

To demonstrate the competing forces of the origin firm and the destination firm effects, I further consider two special cases. Recall that the origin firm effect refers to the upward bias generated because high-paying origin firms tend to be associated with high mover tenure. To verify this source of bias, I let mobility depend only on the current firm's pay level but not the potential firm's pay level:

$$j(i,t) = j^*(i,t) \qquad \iff \beta_1(T_{ij(i,t-1)t-1} + 1) + \eta_{j(i,t-1)} < \nu_{it}$$

The destination firm effect is about the negative correlation between the firm pay level and a worker's tenure since workers on average move toward high-paying firms as they progress in their careers. This effect contributes to a downward bias in the return to tenure estimates. To reflect the selection on the destination firm effect, I let mobility decision depend only on the potential firm pay level:

$$j(i,t) = j^*(i,t) \qquad \iff \beta_1(T_{ij(i,t-1)t-1} + 1) < \eta_{j^*(i,t)} + \nu_{it}$$

Case (4): Mobility based on firm time-invariant and time-varying effects. Now I allow mobility to depend not only on the time-invariant effects of the current and the potential firms but also on their time-varying pay levels.

At the beginning of the first period $t = 1$, each worker i draws his worker type l, firm $j(i,1)$, and an initial tenure level $T_{i,1}$ from a joint distribution F. Independently of everything else, he draws the worker innovation $u_{i,1}$ from $F_u(u)$.

At the beginning of each period t, a worker of type l has probability $p_4(\alpha_l)$ to draw a potential firm $j^*(i,t)$ from $\{1, ..., J\}$. He draws worker innovation u_{it} from $F_u(u)$. Then, he sees the firm-specific wage determinants from both the current firm, $(\eta_{j(i,t-1)}, \tilde{\psi}_{j(i,t-1),t}, T_{ij(i,t-1)t-1})$, and his potential firm $\eta_{j^*(i,t)}, \tilde{\psi}_{j^*(i,t),t}$. The return to spending one extra year at the current firm, β_1, is known to the worker. Next, he makes the mobility decision based on the comparison of the two wage offers. I assume that, conditional on receiving a potential job offer, the mobility function is as follows:

$$m(\alpha_i, \eta_{j(i,t-1)}, \tilde{\psi}_{j(i,t-1),t}, \eta_{j^*(i,t)}, \tilde{\psi}_{j^*(i,t),t}, T_{ij(i,t-1)t-1}, u_{it}|\beta_1)$$
$$\equiv m(\eta_{j(i,t-1)}, \tilde{\psi}_{j(i,t-1),t}, \eta_{j^*(i,t)}, \tilde{\psi}_{j^*(i,t),t}, T_{ij(i,t-1)t-1}|\beta_1)$$

such that

$$\begin{aligned} j(i,t) = j^*(i,t) \quad & \iff \underbrace{\beta_1(T_{ij(i,t-1)t-1}+1) + \alpha_i + \eta_{j(i,t-1)} + \tilde{\psi}_{j(i,t-1),t} + u_{it}}_{\text{current firm}} < \\ & \underbrace{\alpha_i + \eta_{j^*(i,t)} + \tilde{\psi}_{j^*(i,t)t} + u_{it}}_{\text{potential firm}} + \nu_{it} \\ & \iff \beta_1(T_{ij(i,t-1)t-1}+1) + \eta_{j(i,t-1)} + \tilde{\psi}_{j(i,t-1),t} < \\ & \eta_{j^*(i,t)} + \tilde{\psi}_{j^*(i,t)t} + \nu_{it} \end{aligned}$$

To demonstrate the within-firm and the between-firm effects as discussed in the previous section, I consider two additional special cases. Recall that the within-firm effect generates an upward bias since a lower within-firm growth induces a more senior worker to quit. To verify the predicted sign of this source of bias, I let mobility depend on the current firm's time-varying pay component but not the potential firm's:

$$j(i,t) = j^*(i,t) \quad \iff \beta_1(T_{ij(i,t-1)t-1}+1) + \eta_{j(i,t-1)} + \tilde{\psi}_{j(i,t-1),t} < \eta_{j^*(i,t)} + \nu_{it}$$

The between-firm effect is predicted to contribute to a downward bias because a senior worker requires a large increase in the between-firm pay differences as compensation for letting go of the accumulated tenure when switching to the potential firm. To show this between-firm effect, I let mobility decision depend only on the potential firm's time-varying pay component but not the origin firm's:

$$j(i,t) = j^*(i,t) \qquad \Longleftrightarrow \beta_1(T_{ij(i,t-1)t-1} + 1) + \eta_{j(i,t-1)} < \eta_{j^*(i,t)} + \tilde{\psi}_{j^*(i,t),t} + \nu_{it}$$

While the four mobility models described above may be restrictive, it is important to emphasize that the objective of this simulation exercise is to verify how the conjectured mechanisms can generate biases in the predicted and observed directions. Using simple and transparent models of mobility helps deliver clearly how each wage variable plays a role in potentially driving biases in different directions. A more comprehensive model of firm wage-setting and worker sorting should be built and investigated for future research, but it is out of the scope of this paper.

1.7.2 Simulation Results

In the simulation exercise, I choose $N = 10000$, $J = 600$, $n_l = 10$, $n_k = 30$ and $T = 10$. The true return to tenure, β_1, is set to 0.0074 to match the 10-year average of the estimated annual return to job tenure.

The values of α_l are set to the $\frac{l}{n_l} \times 100$-th percentile of the normal distribution with mean zero and standard deviation σ_α. The values of η_k are set to the $\frac{k}{n_k} \times 100$-th percentile of the normal distribution with mean zero and standard deviation σ_η. I assign the estimated standard deviations of α_i and η_j from the FE-TV estimation to σ_α and σ_η, respectively. $\tilde{\psi}_{k(j)t}$ is assumed to be normally distributed IID across $k(j)$ and t with zero mean and standard deviation $\sigma_{\tilde{\psi}}$. $\sigma_{\tilde{\psi}}$ is set equal to the estimated standard deviation. Lastly, u_{it} is assumed to be normally distributed IID across i and t with zero mean and standard deviation σ_u. I let $\sigma_u = 0.005$ to reduce the dispersion of the Monte Carlo distributions.

As to the probabilities of job moves, I set p and $p(\alpha_l)$ to match the empirical shares of movers for each l, where I partition workers into n_l equally sized groups based on the estimated α_i. $p_2(\alpha_l)$, $p_3(\alpha_l)$ and $p_4(\alpha_l)$ are assumed to be a linear combination of $p(\alpha_l)$ and a constant. In the first period, a worker's initial tenure level $T_{i1} \in \{0, 1\}$ is drawn from a binomial distribution with parameter $c(k)$ where $c'(k) > 0$. Modeled this way, a high-paying firm group k is associated with high worker tenure levels in the initial period.

Table 1.10 presents the simulation results. Each panel denotes a mobility model, from the most restrictive (Case (1)) to the least restrictive (Case (4)). Each column corresponds to a wage specification for estimation. In the table, I report the mean of the estimated β_1 net the true value of β_1, averaging across $S = 1000$ simulations.

The first panel shows that all four specifications produce estimates centered around the true value of β_1. This is not surprising, because workers' mobility and the initial tenure are both exogenous to worker types, firm pay levels, and the time-varying firm pay innovations in the data generating process. Consequently, workers' tenure is independent of the remaining variables in the wage model, and the OLS is unbiased.

The data generating process corresponding to the second panel requires that lower type workers have a higher probability of job separation.[17] The heterogeneity in job separation naturally introduces systematic differences in job tenures across worker types. As a result, OLS estimates in the simulation suffer from an upward bias of magnitude 0.0048 (0.0013). Once the worker heterogeneity is controlled for, as in FE-W, FE-TINV, and FE-TV, this upward bias is corrected and the estimates from these three specifications are centered around the true value of β_1.

The third panel considers the mobility function where a worker's mobility decision depends on η's, in addition to α. The first two rows report the estimates where η's of both the current and potential firms enter the mobility decision. When job mobility depends on both firms' η, the OLS and FE-W specifications produce estimates that are biased. In particular, the Monte Carlo distribution of FE-W is centered to the right of the true β_1, with an esti-

17. Empirically, the share of movers falls in l where l denotes worker groups partitioned by quintiles of $\hat{\alpha}_i$.

mated bias of 0.0045 (0.0019). This is the case because, on average, high-paying firms are associated with high tenure levels, and the origin firm effect dominates the destination firm effect in the simulations.

Case (3a) and (3b) in the third panel display the results from the two special cases. Case (3a) considers the case where only the current firm's η enters the mobility function. Confirming the prediction from the previous section, when mobility depends only on the current but not the destination firm, the origin firm effect contributes to a positive bias in the return to tenure. In the simulation, FE-W suffers from an upward bias of roughly 0.0093 (0.0018), compared to FE-TINV. Case (3b) shows the simulation results from the second special case where only the potential firm's η matters to mobility. Verifying the previous prediction, the destination firm effect contributes to a negative bias of -0.0045 (0.0008) in β_1^{FE-W}. In both cases, the Monte Carlo distributions of FE-TINV and FE-TV are centered around the true value.

The last panel in Table 1.10 documents the simulation results where not only firms' η but also their $\tilde{\psi}$'s enter the mobility function. Row 1 and 2 report the estimates when both the current and potential firms' η and $\tilde{\psi}$ are allowed to relate to mobility. The estimated bias in FE-TINV is -0.0042 (0.0016), with the same sign as the empirical bias on the actual data. Moreover, qualitatively the four estimates show a similar pattern to the estimates obtained from the actual data.

In Case (4a), I assume that mobility is independent of the destination firm's $\tilde{\psi}$ to verify if the within-firm effect indeed contributes to a positive bias in $\beta_1^{FE-TINV}$. The simulation results in Table 1.10 verify this prediction, showing an upward bias in $\beta_1^{FE-TINV}$ of 0.0073 (0.0020). The last two rows in Table 1.10 report the estimated bias when mobility is independent of the origin firm's $\tilde{\psi}$. The between-firm effect predicts that workers are more likely to give up a high tenure level if the destination firm has a high draw of $\tilde{\psi}$. As a result, high $\tilde{\psi}$ positively correlates with low tenure level, creating a downward bias in $\beta_1^{FE-TINV}$.

The simulation results are consistent with this prediction. Under this mobility model, the simulated distribution of $\beta_1^{FE-TINV}$ is centered to the left of the true β_1, with a bias of -0.0124 (0.0019). Throughout all cases, the Monte Carlo distributions of FE-TV estimator are centered around the true value.

OLS	FE-W	FE-TINV	FE-TV
Case (1) Random mobility			
0.0001	-0.0001	-0.0001	0.0000
(0.0017)	(0.0023)	(0.0026)	(0.0007)
Case (2) Worker-specific mobility			
0.0048	0.0001	0.0001	0.0000
(0.0013)	(0.0018)	(0.0018)	(0.0000)
Case (3) Mobility based on η's			
0.0114	0.0045	-0.0001	0.0000
(0.0019)	(0.0019)	(0.0019)	(0.0000)
Case (3a) Origin firm effect			
0.0168	0.0093	0.0000	0.0000
(0.0020)	(0.0018)	(0.0017)	(0.0000)
Case (3b) Destination firm effect			
-0.0044	-0.0045	0.0000	0.0000
(0.0013)	(0.0008)	(0.0006)	(0.0000)
Case (4) Mobility based on $(\eta, \tilde{\psi})$'s			
0.0014	-0.0007	-0.0042	0.0000
(0.0018)	(0.0015)	(0.0016)	(0.0000)
Case (4a) Within-firm growth			
0.0113	0.0118	0.0073	0.0000
(0.0022)	(0.0021)	(0.0020)	(0.0000)
Case (4b) Between-firm growth			
-0.0066	-0.0104	-0.0124	0.0000
(0.0017)	(0.0018)	(0.0019)	(0.0000)

Table 1.10: Mean Bias Across Specifications by Mobility Models ($S = 1000$ Replications)

Notes:
The table reports the estimated bias and the standard deviation of the Monte Carlo distribution based on $S = 1000$. $\sigma_\alpha = 0.4, \sigma_\eta = 0.1, \sigma_{\tilde{\psi}} = 0.15, \sigma_u = 0.005$, $\sigma_\nu = 0.2$, and $c(k) = \frac{1}{4} + (5e - 6) \times k$.

1.8 Industry Tenure versus Job Tenure and Worker Observed Heterogeneity

As [49] and [50] emphasize, not all human capital accumulated at the firm level is completely firm-specific. The return to industry-specific human capital accounts for a non-trivial fraction of workers' wage growth, especially for skilled workers ([22]). In this section, I take advantage of the rich information on workers' characteristics and their full working histories in the Norwegian administrative data and extend the discussion of the returns to job tenure by considering the industry tenure. I first assess the relative importance of the return to firm tenures, industry tenures, and general experiences among college male workers. Next, I compare the sources of wage growth for four groups of workers: male workers with college degrees or more, female workers with college degrees or more, male workers with high-school degrees, and female workers with high-school degrees.

1.8.1 Distinguishing Firm Tenures and Industry Tenures

To incorporate the return to industry-specific tenures, I consider the following wage model:

$$w_{it} = \beta_1 T_{ij(i,t)t} + \beta_2 M_{it} + \beta_3 S_{is(i,t)t} + \gamma X_{it} + \alpha_i + \psi_{j(i,t)t} + u_{it} \tag{1.10}$$

where $s(i,t)$ is a function that maps a worker to the corresponding industry in t, and $S_{is(i,t)t}$ measures the worker's experience in industry $s(i,t)$. Similar to Section 3, I impose no parametric assumption on firm tenures, so $T_{ij(i,t)t}$ enters the wage equation as a series of indicators for different firm tenure levels. Both industry tenures $S_{is(i,t)t}$ and general labor market experiences M_{it} are modeled as 6th-degree polynomials in the wage equation.

Identification assumption. Since the assumption in equation (1.4) imposes that the worker-specific wage innovation, u_{it}, is conditionally mean independent of the worker's choices of firms, it implies that u_{it} is also conditionally mean independent of the worker's choices of industries and the duration to work in each industry. That is, the assumption in equation (1.4) implies that

$$E[u_{it}|\mathbf{T}, \mathbf{S}, \mathbf{M}, \mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\psi}] = 0$$

where $\mathbf{T} \equiv \{T_{ij(i,t)t}\}_{i\in\{1,...,N\},t\in\{1,...,T\}}$, $\mathbf{S} \equiv \{S_{is(i,t)t}\}_{i\in\{1,...,N\},t\in\{1,...,T\}}$, $\mathbf{M} \equiv \{M_{it}\}_{i\in\{1,...,N\},t\in\{1,...,T\}}$, $\mathbf{X} \equiv \{X_{it}\}_{i\in\{1,...,N\},t\in\{1,...,T\}}$, $\boldsymbol{\alpha} \equiv \{\alpha_i\}_{i\in\{1,...,N\}}$, and $\boldsymbol{\psi} \equiv \{\psi_{jt}\}_{j\in\{1,...,T\},t\in\{1,...,T\}}$.

Estimation sample. In order to measure a worker's industry experiences, I impose additional sampling criteria. In particular, I restrict to the set of workers between 1995 and 2018 who (i) are at least 15 years old in the year 1995, and (ii) who completed their highest degree of education no earlier than 1995.[18][19] With the complete working histories of the individuals in this sub-sample, I then compute industry tenures for each worker in each year.

1.8.2 Return to Firm Tenure versus Return to Industry Tenure

I begin by showing how much the estimated return to firm tenure changes due to the inclusion of industry tenures in the college male sample. In Figure 1.15, the dotted line presents the estimated return to firm tenure without controlling for industry tenures, and the solid line shows the estimates with industry tenures. The return to firm tenures falls by around 1.2 log points in the first two years once the industry tenures are taken into consideration. In the human capital framework, this suggests that a small but non-trivial fraction of human capital accumulated at the firm level is portable within industries. With industry-specific

18. The age criterion is based on the compulsory schooling in Norway.

19. I define a worker's entry year to the labor force as the year when the worker finishes his highest degree of education recorded by 2018.

human capital, workers can switch firms within the same industries without losing all human capital accumulated at previous firms.

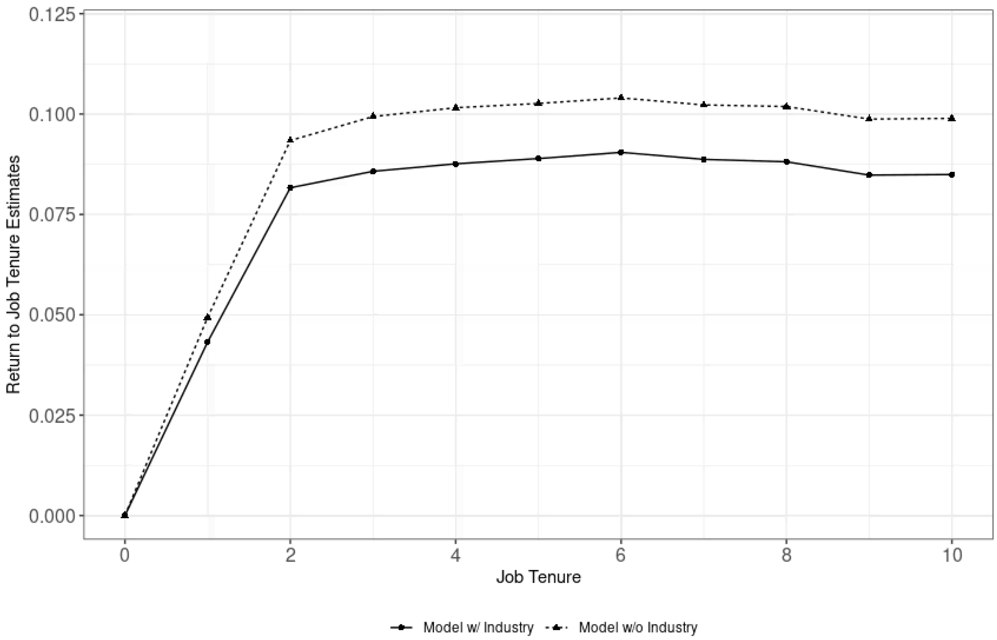


Figure 1.15: Return to Firm Tenure in Industry Model

Notes:
[1] This figure compares the return to firm tenure estimates between two model specifications. The solid line standards for the model where industry tenures are included. The dotted line represents the estimates from the model with no industry tenures.
[2] The estimates are obtained by estimating the wage model (1.10) on college male workers. The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. Both the industry tenure $S_{is(i,t)t}$ and the experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[3] All estimates are obtained from a sub-sample of workers who (i) are at least 15 years old in the year 1995, and (ii) completed their highest degree of education no earlier than 1995.

Next, I compare the relative magnitudes of the return to firm tenures, the return to industry tenures, and the return to general labor market experiences, for the first 10 years since the labor market entrance for college male workers in Norway. Figure 1.16 displays the estimated wage profiles, and Table 1.11 summarizes the estimation results. For young male workers with college degrees in Norway, the accumulation of the general labor market experiences is the driving force for wage growth, accounting for more than 50% of wage growth in the first ten years. The wage growth from firm tenures is substantial in the first two years at a firm, with an average of 4.1 percent per year. However, this wage growth from the return to firm tenures falls back to essentially zero in the subsequent years. In comparison, the return to labor market experiences remains a driving force in these workers' wage growth. The estimated return to industry tenures is much smaller but non-trivial throughout the first 10 years of the workers' career, accounting for 14.0% of wage growth in the first 10 years of these workers' careers.

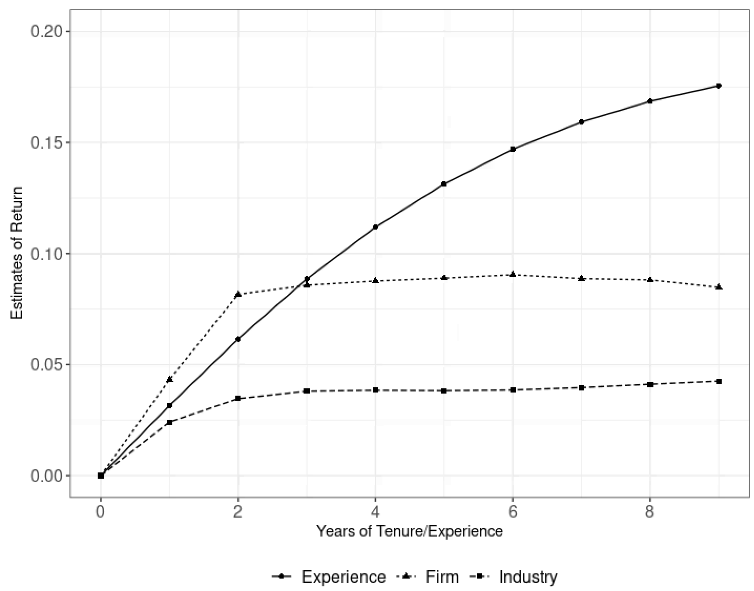


Figure 1.16: Returns to Firm Tenure, Industry Tenure, and Experiences Among College Male Workers

Notes:
[1] The figure shows the returns to firm tenure, industry tenure and experiences during the first ten years for college male workers.
[2] College male workers are defined as male workers whose highest degree of education observed by 2018 is college or above.
[3] The estimates are obtained by estimating the wage model (1.10). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. Both the industry tenure $S_{is(i,t)t}$ and the experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from a sub-sample of workers who (i) are at least 15 years old in the year 1995, and (ii) completed their highest degree of education no earlier than 1995.

1.8.3 Worker Observed Heterogeneity

In this subsection, I assess the relative importance of the return to job tenures, the return to industry tenures, and the return to general labor market experiences across four subgroups of workers: male and female workers with college degrees or above, and male and females workers with high-school degrees.

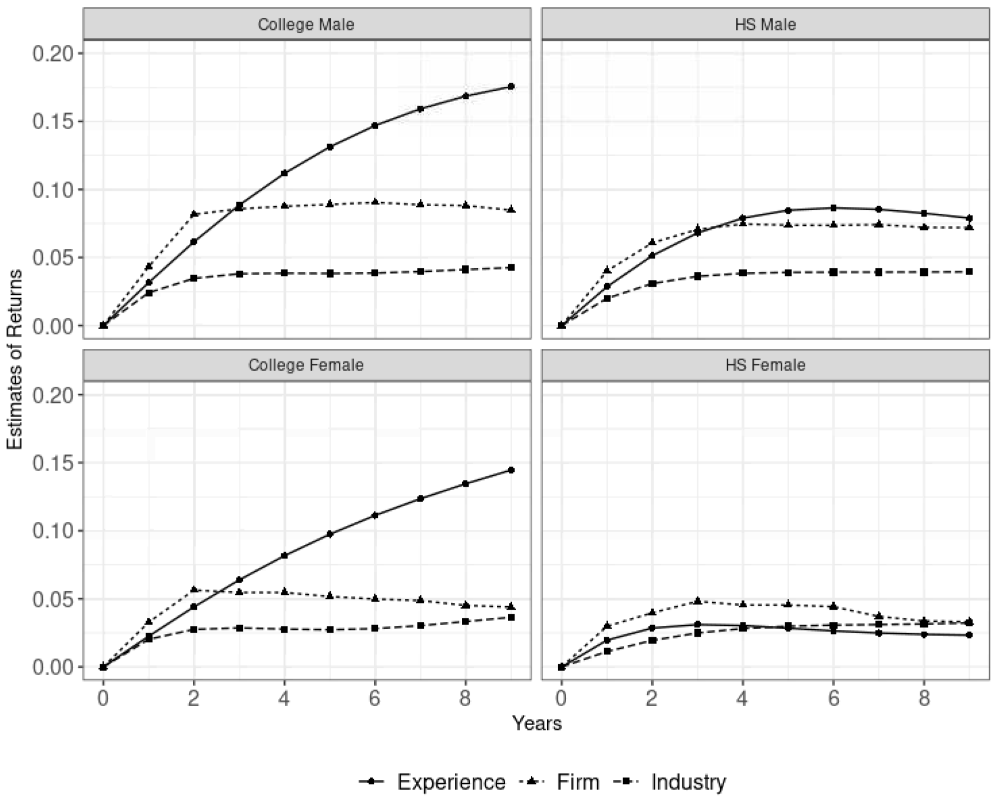


Figure 1.17: Returns to Firm Tenure, Industry Tenure and Experiences Across Worker Groups

Notes:
[1] This plot shows the estimated returns to firm tenure, industry tenure and experiences during the first 10 years across four groups of workers: college male, college female, HS male and HS female.
[2] College male (female) is defined as workers whose highest degree of education observed by 2018 is college or above. HS male (female) is defined as workers whose highest degree of education observed by 2018 is high-school degree.
[3] The estimates are obtained by estimating the wage model (1.10). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. Both the industry tenure $S_{is(i,t)t}$ and the experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from a sub-sample of workers who (i) are at least 15 years old in the year 1995, and (ii) completed their highest degree of education no earlier than 1995.

Results

In Figure 1.17, I plot the profiles of the three sources of wage growth across the four worker groups. Table 1.11 summarizes the results. Similar to college male workers, the primary and lasting source of wage growth for college female workers is the accumulation of general labor market experiences. On average, wage growth from labor market experiences is around 1.54 log points per year in the first ten years since the workers' entrance into the labor market. The wage growth from the return to firm tenure in the first two years is substantial but smaller than the return enjoyed by college male workers, averaging around 2.82 log points per year. The return to the industry tenures is about half the size of the return to firm tenures.

Workers with high school degrees see a much different return-to-experience profile as compared to those with college degrees. In particular, while the return to general labor market experiences is as high as 1.69 log points per year for high school male workers (0.57 log points for high school female workers) in the first 5 years, the return declines to effectively 0 afterward. Moreover, among these workers, the magnitude of the return to the general labor market experiences is comparable to the return to firm tenures. For male workers with high school degrees, industry return is about half the size of the return to firm tenures. For female workers with high school degrees, the magnitudes of all three returns are roughly the same and much lower compared to other groups of workers.

	College Male	College Female	HS Male	HS Female
2 Years of Firm Tenure	0.0816	0.0564	0.0607	0.0397
	(0.0010)	(0.0014)	(0.0013)	(0.0031)
2 Years of Industry Tenure	0.0346	0.0275	0.0309	0.0195
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
2 Years of Experiences	0.0615	0.0441	0.0514	0.0285
	(0.0000)	(0.0000)	(0.0001)	(0.0002)
5 Years of Firm Tenure	0.0889	0.0517	0.0737	0.0455
	(0.0013)	(0.0019)	(0.0017)	(0.0043)
5 Years of Industry Tenure	0.0382	0.0273	0.0391	0.0299
	(0.0001)	(0.0002)	(0.0001)	(0.0001)
5 Years of Experiences	0.1313	0.0975	0.0847	0.0284
	(0.0000)	(0.0000)	(0.0001)	(0.0003)
10 Years of Firm Tenure	0.0849	0.0440	0.0712	0.0355
	(0.0020)	(0.0030)	(0.0024)	(0.0069)
10 Years of Industry Tenure	0.0433	0.0389	0.0393	0.0335
	(0.0002)	(0.0004)	(0.0002)	(0.0001)
10 Years of Experiences	0.1807	0.1541	0.0753	0.0228
	(0.0000)	(0.0000)	(0.0001)	(0.0006)
Mean log wages	3.856	3.626	3.594	3.322

Table 1.11: Firm Tenures, Industry Tenures and Experiences Across Worker Groups

Notes:
[1] This table reports the estimates of the returns to firm tenure, industry tenure and experiences across four groups of workers: college male, college female, HS male and HS female.
[2] College male (female) is defined as workers whose highest degree of education observed by 2018 is college or above. HS male (female) is defined as workers whose highest degree of education observed by 2018 is high-school degree.
[3] The estimates are obtained by estimating the wage model (1.10). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. Both the industry tenure $S_{is(i,t)t}$ and the experience variables M_{it} enter the wage equation as 6th degree polynomials. X_{it} consists of marital status, and occupation dummies.
[4] All estimates are obtained from a sub-sample of workers who (i) are at least 15 years old in the year 1995, and (ii) completed their highest degree of education no earlier than 1995.

The distinct returns to labor market experiences are the main contributing factor to the gap in the wage growth rate between college and high school graduates, especially in the later stage of workers' careers. Within the same gender group, high school graduates see similar but slightly smaller returns to firm and industry tenure as compared with college graduates. In the human capital framework, the results suggest that both firm-specific and industry-specific human capital accumulation rates are similar across workers from different educational backgrounds. However, the highly educated workers are better at accumulating general human capital over time. Female workers generally receive smaller returns from all 3 dimensions as compared to male workers. The differences in the returns to job tenures and experiences contribute the most to the wage growth gap between male and female workers in the first 10 years since the labor market entry.

1.9 Conclusion

In this paper, I examine how and to what extent the selection of job moves on firm pay innovations resulting from the worker-firm rent-sharing responses generate biases in the return to tenure estimators. I show empirically that this downward bias arises because shocks to firms' productivity induce firms to bid up wages to increase hiring, creating a positive relationship between movers' foregone tenure levels and the destination firm's shocks. I embed [1] framework into a Mincerian-style wage model and extend it by further allowing for flexible time-varying firm fixed effects. The estimation results on the matched employer-employee data in Norway reveal that the downward bias is large. I derive and decompose the bias into two pieces that are related to firm pay innovations at the origin and the destination firms. I show that the estimated firm time-varying effects are positively correlated with but not sufficiently explained by firm-level productivity measures. Alternative external measures such as the share of newly hired workers instead help reduce the bias in the return to job tenure estimator.

Future research is called for to better understand the determinants of the time-varying firm pay components and the reasons behind the poor performance of firm-level productivity measures in reducing the bias. Moreover, unobserved interaction effects between workers and firms are likely to further create biases in the return to tenure estimates. Worker-specific wage innovations that are not common within firms but correlate with mobility can also introduce potential biases. Future work can focus on taking into account these two additional sources of unobserved heterogeneity in the study of the return to tenure estimation.

Works on identifying and quantifying the dispersion in firm-specific return to job tenure is another extension to this paper. As is stressed by this paper, accounting for the time-varying firm pay policies in workers' wages is of great importance to mitigate biases in the return to tenure estimators. Therefore, developing a model with reliable identification and estimation strategies that separates firm-specific wage profiles from firm time-varying effects is the key to answering the above question.

1..1 Additional Figures

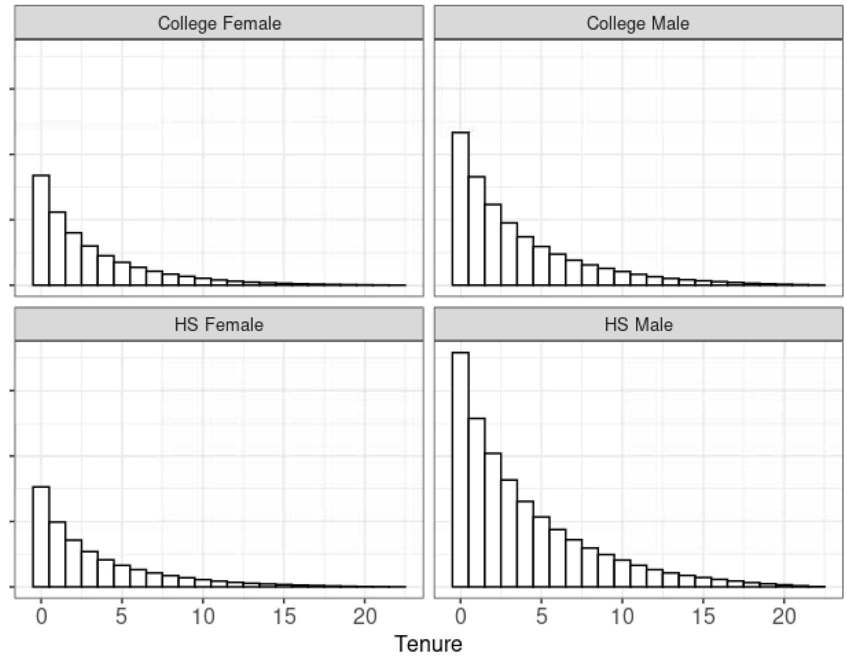


Figure 1.18: Distribution of Tenure by Worker Groups

Notes:
[1] This figure plots the histogram of tenures across four groups of workers in the private sector: college male, college female, HS male and HS female.
[2] College male (female) is defined as workers whose highest degree of education observed by 2018 is college or above. HS male (female) is defined as workers whose highest degree of education observed by 2018 is high-school degree.

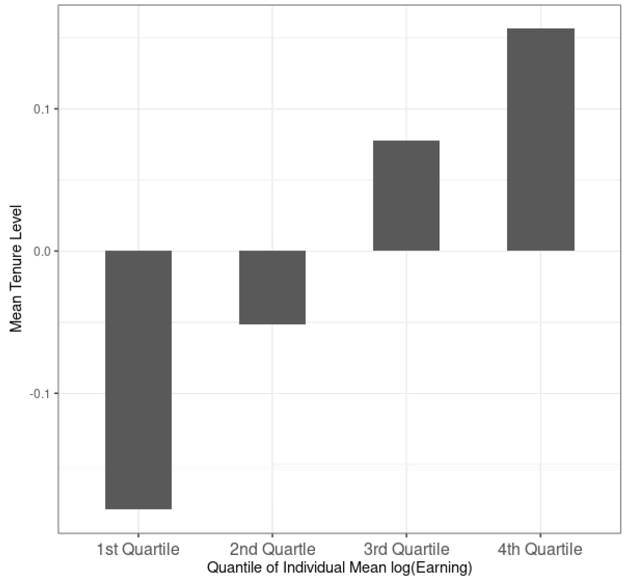


Figure 1.19: Covariance between Individual Mean Wage and Residualized Tenure

Notes:
[1] This figure plots the average residualized tenure level across 4 groups of workers ranked by their individual-specific mean wages. To construct this plot, I first residualize workers' wages by aggregate market $\times$year effects as well as worker characteristics. I then use the residualized wages to construct individual-specific mean wages for each worker in the sample. Next, I rank individuals based on their mean wages and partition individuals into 4 quartile groups. The y-axis of this plot is the residuals from regressing a linear tenure variable on workers' experiences, market $\times$ year effects, and other worker characteristics.
[2] The result is based on the largest connected set of college male workers in the private sector. The connected set is defined by a network of workers and firms.

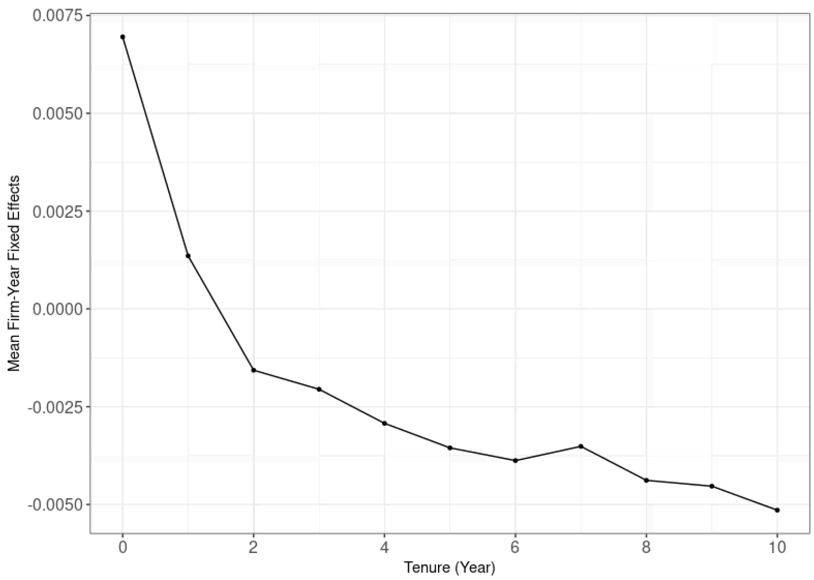


Figure 1.20: $\tilde{\psi}_{jt}$ over Tenure

Notes:
The figure shows the estimated $\tilde{\psi}_{jt}$ across job tenures. A clear inverse relationship between job tenures and time-varying firm fixed effects confirms the negative bias in the return to tenure estimates. The y-axis is the idiosyncratic component in the estimated firm pay premiums, $\tilde{\psi}_{jt}$, rather than ψ_{jt}.

1..2 Firm-Level Observables and Biases

Var. Raw $\hat{\psi}$	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)
Log(VA)	0.0183						0.0070	0.0050
	(0.0001)						(0.0003)	(0.0004)
Lagged log(VA)		0.0133						0.0023
		(0.0001)						(0.0004)
Log(sales)			0.0142				0.0021	0.0027
			(0.0001)				(0.0004)	(0.0005)
Log(revenue)				0.0055			0.0027	0.0013
				(0.0001)			(0.0001)	(0.0001)
Log(firm size)					0.0449		0.0325	0.0415
					(0.0003)		(0.0004)	(0.0006)
Lagged log(firm size)						0.0175		-0.0186
						(0.0002)		(0.0006)
R^2	0.816	0.832	0.814	0.830	0.810	0.813	0.834	0.864
Obs.	1,745,011	1,645,492	1,902,336	1,578,383	2,240,704	2,195,416	1,405,588	1,120,119
Firm FE	x	x	x	x	x	x	x	x
Mkt×yr	x	x	x	x	x	x	x	x
Weights	x	x	x	x	x	x	x	x

Table 1.12: Firm Observables and ψ_{jt} with Firm FE on Unbalanced Sample, Weighted by Employment Size

Notes:
[1] This table reports the estimates of regressing the estimated $\hat{\psi}_{jt}$ on the firm-level variables.
[2] The variable "VA" refers to value-added, "revenue" refers to operating revenues.
[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation so that $\hat{\psi}_{jt}$ are observed. Since value-added, sales and operating revenues are not observed or are negative for some observations, the estimation samples vary across Column (1) to (9).
[4] Firm employment sizes are used as weights in the estimation.

Var. $\hat{\psi}$	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Tenure	-0.0011	-0.0008	-0.0007	-0.0011	-0.0010	-0.0006	-0.0004	-0.0004	-0.0003	-0.0003
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Log(VA)		0.0194	0.0159			0.0119	0.0119	0.0080	0.0072	0.0071
		(0.0002)	(0.0002)			(0.0002)	(0.0002)	(0.0003)	(0.0005)	(0.0005)
Lagged Log(VA)			0.0093			0.0077	0.0087	0.0071	0.0042	0.0042
			(0.0002)			(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0004)
Log(size)				0.0476	0.0618	0.0510	0.0289	0.0287	0.0241	0.0289
				(0.00003)	(0.0004)	(0.0005)	(0.0007)	(0.0007)	(0.0008)	(0.0011)
Lagged Log(size)					-0.0181	-0.0217	-0.0036	-0.0039	-0.0029	-0.0058
					(0.0003)	(0.0004)	(0.0006)	(0.0006)	(0.0007)	(0.0010)
New-hire share)							0.0564	0.0562	0.0510	0.0452
							(0.0013)	(0.0013)	(0.0016)	(0.0020)
Lagged New-hire share							0.0298	0.0299	0.0263	0.0263
							(0.0007)	(0.0007)	(0.0008)	(0.0008)
Log(sales)								0.0056	0.0028	0.0026
								(0.0004)	(0.0005)	(0.0005)
Lagged Log(sales)								0.0020	0.0010	0.0010
								(0.0004)	(0.0001)	(0.0005)
Log(revenue)									0.0014	0.0014
									(0.0001)	(0.0001)
Lagged Log(revenue)									0.0042	0.0042
									(0.0001)	(0.0001)
Quit ratio										-0.0194
										(0.0015)
Lagged Quit ratio										0.0075
										(0.0019)
Mkt×year	x	x	x	x	x	x	x	x	x	x
Firm FE	x	x	x	x	x	x	x	x	x	x
Worker X	x	x	x	x	x	x	x	x	x	x
R^2	0.854	0.865	0.875	0.862	0.876	0.876	0.0.876	0.899	0.899	
Obs.	2,240,704	1,745,011	1,595,534	2,240,704	2,195,416	1,595,389	1,595,389	1,595,071	1,157,737	1,157,737

Table 1.13: ψ_{jt} and Tenure with Firm Observables, Unbalanced Sample

Notes:

[1] This table reports the estimates of regressing $\hat{\psi}_{jt}$ on a linear tenure term and the firm-level variables.

[2] The variable "VA" refers to value-added, "revenue" refers to operating revenues, "size" refers to firm sizes, "new-hire share" refers to the employment share of newly hired workers, and "quit ratio" refers to the employment share of quitters. Worker X refers to worker-level controls including marital status, occupation dummies and experiences (6th degree polynomials).

[3] The sample of the largest connected set of workers and firm×years among college male workers is used for estimation so that $\hat{\psi}_{jt}$ are observed. Since value-added, sales and operating revenues are not observed or are negative for some observations, the estimation samples vary across Column (1) to (9).

[4] Firm employment sizes are used as weights in the estimation.

	FE-TINV	FE-TV	Value-Added	Firm Size	Value-Added Firm Size	Share of New Hires	Share of New Hires and Quitters	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2 Years of Tenure	0.0899	0.0831	0.0824	0.0851	0.0850	0.0890	0.0889	0.0876
	(0.0007)	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
5 Years of Tenure	0.0911	0.0814	0.0806	0.0829	0.0830	0.0892	0.0893	0.0871
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
10 Years of Tenure	0.0782	0.0676	0.0678	0.0690	0.0698	0.0748	0.0748	0.0738
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
$\log(VA_{jt})$			0.0146		0.0116			0.0113
			(0.0003)		(0.0003)			(0.0003)
$\log(VA_{jt-1})$			0.0105		0.0091			0.0093
			(0.0003)		(0.0003)			(0.0003)
$\log(VA_{jt-2})$			-0.0006		-0.0008			-0.0003
			(0.0003)		(0.0003)			(0.0003)
$\log(size_{jt})$				0.0567	0.0448			0.0201
				(0.0009)	(0.0009)			(0.0012)
$\log(size_{jt-1})$				-0.0141	-0.0171			-0.0037
				(0.0008)	(0.0008)			(0.0010)
$\log(size_{jt-2})$				-0.0044	-0.0060			0.0047
				(0.0006)	(0.0006)			(0.0008)
New hire share in t						0.0802	0.0863	0.0707
						(0.0017)	(0.0017)	(0.0022)
New hire share in $t-1$						0.0386	0.0412	0.0360
						(0.0015)	(0.0015)	(0.0019)
New hire share in $t-2$						0.0115	0.0109	0.0096
						(0.0012)	(0.0012)	(0.0012)
Quitter share in t							-0.0277	
							(0.0018)	
Quitter share in $t-1$							-0.0335	
							(0.0019)	
Quitter share in $t-2$							-0.0162	
							(0.0018)	

Table 1.14: Return to Tenure with Firm-Level Observables

Notes:
[1] This table reports the estimates of the return to job tenure using the wage model (1.9). The tenure term $T_{ij(i,t)t}$ enters the model non-parametrically. The experience variable M_{it} enters the wage equation as a 6th degree polynomial. X_{it} consists of marital status, and occupation dummies.
[2] $Z_{j(i,t)}$ is a set of firm-level variables. For Panel "Value-Added", Z_{jt} includes the log value-added in period t, $t-1$ and t. For Panel "Firm Size", Z_{jt} includes the log firm size in period t, $t-1$ and t. For Panel "Value-Added and Firm Size", Z_{jt} includes the log value-added and log firm size in period t, $t-1$ and t. For Panel "Share of New Hires", Z_{jt} includes the number of newly hired workers divided by the employment size in period t, $t-1$ and t. For Panel "Share of New Hires and Quitters", Z_{jt} includes the share of new hires and the number of quitters/layoffs divided by the employment size ain period t, $t-1$ and t. For Panel "Full", Z_{jt} includes the log value-added, log firm size, the share of new hires and the share of quitters in period t, $t-1$ and t.
[3] The red line represents the FE-TINV estimates. The black line represents the FE-TV estimates. The blue line represents the estimates from model (1.9) with Z_{jt} corresponding to each panel.
[4] All estimates are obtained from the largest connected set of worker and firm×year where all variables in Z_{jt} are observed (and log value-added is greater than 0).

CHAPTER 2

RESERVATION WAGES AND WORKERS' VALUATION OF JOB FLEXIBILITY: EVIDENCE FROM A NATURAL FIELD EXPERIMENT

2.1 Introduction

The last decade has witnessed a dramatic increase in the prevalence of flexible working, either via workers entering the Gig Economy or historically traditional jobs becoming more flexible, allowing the worker to choose specific hours or where to work. These changes raise several questions of both policy and practical importance. How do labor supply elasticities and reservation wages vary across days of the week and hours of the day? To what extent do labor supply elasticities and reservation wages differ between people such as men and women or old and young? How do different workers value the ability to customize work schedules? While both economists and policymakers are keenly interested in these questions, credible answers have been hindered by a lack of high frequency panel data on wages and work decisions as well as by the difficulty of identifying how labor supply elasticities, reservation wages, and the value of flexibility vary between people and over time.

The goal of our paper is to answer the above questions while addressing both the measurement and the identification challenges. The context of our study is the largest ride-sharing company, Uber. Our work draws on three strengths of this environment.[1] First, Uber allows a driver to work anytime she is willing to accept the wage she would be paid in the market. Second, we have access to high frequency panel data on the wage an individual is paid and her decision to work.[2] Third, via a large natural field experiment we observe reactions to exogenous variation in expected market wages across individuals and over time.

Combining the panel data with the experiment, we first estimate individuals' labor supply responses to exogenous changes in expected market wages. These experimental findings motivate and guide our modeling of the labor supply of the drivers. The primitives of the

1. [29] describe the labor market for Uber's drivers. They find that drivers cite flexibility as a reason for working for Uber and that many drivers report that Uber is a part-time activity secondary to more traditional employment.

2. As in [18] and [14], we calculate the "wage" in an hour as a driver's total earnings in that hour divided by minutes worked (i.e. the number of minutes for which a driver has the app on and is available for accepting requests).

model are recovered from a combination of the experimental estimates and other data moments. We use the estimated model to compute how labor supply elasticities and reservation wages vary between people and over time and to perform counterfactual analyses. These analyses allow us to infer the drivers' willingness to pay for flexible work arrangements.[3]

In Section **??**, we describe the labor market for Uber's drivers and the natural field experiment. The analyses of the experiment yield three main findings, which we present in Section **??**. The first main finding is that the labor supply responses vary systematically both across people and over time. In order to discover these heterogeneous effects, we apply the method of [15] to estimate an instrumental variables model with a full set of interactions between the endogenous regressor of interest, wages, and the pre-determined covariates, gender, hours of the day, and days of the week. A clear pattern of heterogeneity emerges: Labor supply is most responsive in the evenings; men and older drivers have, on average, larger responses than other drivers. Taken together, these results suggest it is key to allow preferences of the drivers to vary by gender and age for each demographic group across hours of the day and days of the week.

The second main finding is that drivers do not only increase their labor supply during the periods with exogenously higher wages but also in the hours preceding and following these periods. This finding of anticipatory and persistent responses to increases in expected wages is consistent with forward looking drivers with fixed costs of starting to drive. In the presence of such adjustment costs, a static labor supply model is insufficient to analyze the behavior of the drivers. A dynamic model is needed to capture the connection between the decision to drive in the current period and future utility.

The third main finding is that unobserved determinants of wages, if ignored, lead to a significant downward bias in the estimated labor supply responses. In particular, OLS esti-

3. In the working paper version of this paper (https://www.nber.org/papers/w27807), we also perform another counterfactual experiment which allows us to examine how preference heterogeneity and adjustment costs influence the effectiveness of driver incentives that Uber may offer.

mates show much weaker associations between labor supply and wages than the experimental estimates. This downward bias is consistent with demand being high when it is costly for the drivers to work. Including fixed effects for workers, days of the week, and hours of the day reduces the bias, but the labor supply elasticities remain too small. This finding suggests that idiosyncratic factors, such as weather conditions and entertainment events, may create high demand while, at the same time, make driving more costly or difficult.

The experimental estimates provide key data points for recovering reservation wages, labor supply elasticities and the value of job flexibility, but do not by themselves tell us these quantities. To do so, we develop, in Section 4, a dynamic model of labor supply. This model builds on the experimental findings and accommodates important features of the market, including uncertainty about wages and costs of driving in the future, the possibility of a job other than driving for Uber, and fixed costs of starting to drive. When taking the model to the data, we allow for both observed and unobserved heterogeneity across drivers, and we allow market wages to be correlated with the cost of driving in a given period. Even with these considerations, it is possible to prove identification of the primitives of the model given the panel data and the experiment that creates exogenous variation in expected market wages.

We use the EM algorithm to find the maximum likelihood estimates of the model parameters. The parameter estimates suggest significant costs of starting to drive and considerable observed and unobserved heterogeneity in preferences. Conditional on age and gender, there appears to be three types of drivers: The 'infrequent driver' who only drives occasionally; the 'full-time driver' who drives regularly both in the evening and during the day; and 'the evening driver' who rarely drives during the day, possibly because she has a daytime job. To assess the importance of substitution between Lyft and Uber, we compare the results for all drivers to those we obtain from a subsample of drivers who are ineligible to drive for Lyft. It is reassuring to find that both the experimental estimates and the estimated model parameters do not materially change when we restrict attention to this subsample.

The model delivers two key insights, presented in Section 5. The first insight from the model is that reservation wages, and thus the shadow prices of time, vary a lot both over time and across people.[4] For an average driver at a typical day, the reservation wage is relatively low during the day, starts increasing in the late evening, peaks at around 4 a.m., and then declines gradually until 9 a.m. By way of comparison, there is little variation in reservation wages across days of the week: On average, the reservation wage is only a few percent lower on weekdays than during the weekends. Holding day of the week and hour of the day fixed, there is also a great deal of variation in reservation wages across people. On average, reservation wages of women are 106 percent higher than male reservation wages. There is also a great deal of heterogeneity conditional on observables: The infrequent drivers have much higher reservation wages than the full-time drivers, while the evening drivers demand relatively high wages to drive during the day.

The second insight from the model is that drivers would demand much higher wages if they had to commit to pre-set work schedules. We quantify the importance of two distinct types of job flexibility. One is the ability to set a customized work schedule, so that each driver may plan to work only when her expected reservation wage is lower than the expected wages. We quantify the value of this type of flexibility by removing certain hours of the day or days of the week from the choice set of the workers. Our findings suggest that drivers are particularly averse to restrictions on what hours of the day to work. By way of comparison, constraining drivers to work only on the weekends or only on weekdays would require a modest increase in wages. The other type of flexibility we consider is the ability to adjust the schedule from day to day or even hour to hour in response to unexpected changes to offered wages or costs of driving. We measure the value of this flexibility by

4. It is important to recognize that a driver's reservation wage should be interpreted as a shadow price of time that reflects not only leisure possibilities but also alternative economic activities such as home production or other jobs.

restricting drivers to stick to the work schedule they prefer before observing any shocks to wages and preferences. Our findings suggest that Uber drivers, especially those who are female, benefit significantly from the possibility to adapt work schedules to unexpected events. Taken together, these results suggest that job flexibility is a central component of the total compensation of ride-sharing companies like Uber.

Our paper is primarily related to a large literature on labor supply. The models, data, and findings have been summarized and critiqued in multiple review articles including [51], [39], [9], [38] and [16]. Most models of labor supply are concerned with the problem of choosing how much to work, not when to work. In many of these models, there are no hours restrictions, and each worker supplies labor until the wage she would face in the market equals the value she places on her time, the reservation wage. When taking such models to the data, labor supply elasticities and reservation wages are typically inferred from differences in work hours across people given their observed wages. There are, however, several concerns with this revealed preference argument. One of these concerns is that both theory and evidence suggest restrictions on hours choices stemming from the demand side of the market. This concern motivates a large body of work that incorporates hours restrictions in models of labor supply under the assumption that the analyst has full or partial knowledge about the probability distributions of either offered or desired hours of work.[5]

To avoid making questionable assumptions about hours restrictions, we take advantage of the fact that Uber is a platform on which drivers, once approved, are free to choose their work hours. There are no minimum-hours requirements and only modest constraints on maximum hours. As a result, our estimates of extensive margin labor supply elasticities and reservation wages are not confounded by hours restrictions from the demand side of the market. Instead, the estimated elasticities capture the sensitivity of the decision to supply labor in a given hour to anticipated and exogenous changes in hourly market wages. The wage changes we

5. See [8] and the references therein for details.

consider are modest and temporary, so that lifetime wealth is approximately unchanged. Thus, our setting allows us to recover estimates of extensive margin Frisch elasticities per hour and elasticities of intertemporal substitution (IES) between hours, which in our model differ due to adjustment costs.

Averaging over time and across drivers, we find an extensive margin Frisch elasticity of 0.65, and an IES of 0.45. The estimated Frisch elasticity is significantly larger than what is typically reported in micro studies that ignore or make assumptions about hours restrictions from the demand side of the market ([16]). Our IES falls in the range of 0.22 and 0.60, comparable to those by [25] and [26]. By contrast, [5] report estimates of IES close to one. Their estimates are based on a comparison of the commission-based compensation model of Uber and the conventional taxi contract. However, as emphasized by [46] in their review of the literature, it is difficult to compare the estimates of IES across studies, in part because the restriction on hours may vary but also because the accounting period differs (e.g. days, weeks, or years).

The closest study to ours is arguably the work of [14]. Like us, they take advantage of the fact that Uber has virtually no hours restrictions.[6] Thus, [14] argue, one can recover how reservation wages vary across people and time by relating the probability an individual drives in a given time period to the mean prevailing market wage for that period. Using a multivariate probit model with time-varying thresholds for work decisions, they estimate driver-specific reservation wages, and then decompose these reservation wages into predictable and unpredictable components. Armed with the estimates from this static labor supply model, they calculate the surplus from driving for Uber and the surplus changes that would result from requiring the driver to instead work specific patterns of hours.

Our paper complements and extends the model and analyses of [14] in several important

6. There are also other papers using data from Uber. [5] study how workers' view the commission-based compensation model of Uber as compared to traditional taxi compensation contract. [17] estimate consumers' demand and surplus from Uber rides. [18] study the determinants of the gender earnings gap amongst Uber drivers.

ways. First, motivated by our natural field experimental results, we develop, identify, and estimate a dynamic model of labor supply with fixed costs of starting a shift. Second, we allow market wages to be correlated with the unobserved cost of driving in a given period. Third, we use a natural field experiment to identify the primitives of the model. Fourth, we allow permanent heterogeneity both by the drivers' observable characteristics and according to their unobserved latent types. Empirically, we find that these modeling choices are important to match the data as well as for the estimated reservation wages and the counterfactual analyses. Our paper also offers a complementary perspective on the heterogeneity in reservation wages. [14] model and estimate the heterogeneity in reservation wages as arising from idiosyncratic preferences. We show there is a systematic and predictable pattern in the reservation wages by not only the day of week or hour of the day but also according to the gender, age, and type of driver. This pattern is useful to better understand who benefits from flexible work arrangements, and, as a result, it may also help improve the design of driver incentives and inform discussions over recent policy proposals about regulation and pay rules for ride-sharing companies.

Our paper also relates to a body of work on the labor supply of taxi drivers. The primary goal of this work is to estimate the wage elasticity of daily hours of work to test if labor supply behavior is consistent with reference dependence. The work is summarized and critiqued in [26]. He also replicates and extends existing work. His findings suggest that reference dependence is not an important factor in the daily labor supply decisions of taxi drivers.

Some of our findings are similar to those reported in [26]. For example, much of the variation in hourly wages is predictable based on the day of the week and the hour of the day, and drivers are more likely to work when market wages are high. Other findings differ. For instance, [26] finds that the probability of ending a shift depends strongly on hours worked. We do not find support for such fatigue being empirically important for the behavior of Uber drivers. However, the environment and decision problem of Uber drivers differ in important ways as compared to taxi drivers. In particular, accumulated hours worked in a given day tend to be a lot higher for taxi drivers, and, as a result, fatigue could be more salient for whether they continue driving or end a shift. By comparison, the labor supply of Uber drivers is best described by a combination of adjustment costs in terms of starting to drive and heterogenous reservation wages, especially by hour of the day and type of driver.

Another literature to which we relate is the research on how individuals value workplace amenities such as job flexibility. Survey evidence shows that workers state that they are willing to take lower pay for more flexible jobs (e.g., [31]; [52]; [24]; [44]). However, recovering the workers' actual valuation of job flexibility from naturally occurring data has proven difficult for several reasons. One challenge is that firms may pay differently simply because they employ workers of different quality. A second challenge is that observed wage variation across firms may reflect workplace amenities other than job flexibility. Most research to date tries to address these issues by controlling for worker and firm characteristics, hoping that any remaining wage variation across firms is due to job flexibility.[7]

Even if these controls were sufficient to address concerns about omitted variables bias, it is important to observe that additional assumptions or data are needed to draw inference about workers' valuation of job flexibility. Wage differentials across firms could reflect imperfect competition in the labor market, not workplace amenities. Additionally, in standard models of equalizing differences, such as [53], the observed wage differentials are the market prices of amenities, providing only information on the valuation of marginal workers. [**?**] develop, identify and estimate an equilibrium model of the U.S. labor market with two-sided heterogeneity where workers view firms as imperfect substitutes because of heterogeneous preferences over workplace amenities. The estimated model makes it possible to distinguish between and draw inference about imperfect competition, compensating differentials, and

7. [47], [14] and [33] review this literature.

the distribution of worker preferences over amenities. The empirical findings suggest one needs to be cautious in extrapolating the valuation of amenities among marginal workers, as measured by the compensating differentials, to the valuation of inframarginal workers, who extract a significant amount of surplus or rents from workplace amenities. The importance of worker heterogeneity in the value of amenities like job flexibility is consistent with both our findings and those in [14].

Our findings on job flexibility complement recent evidence that uses a stated preference approach to infer workers' preferences based on their choices between pairs of exogenously assigned hypothetical jobs with different combinations of amenity levels and pay (e.g., [47]; [58]). For example, [47] use a discrete choice experiment in hiring for a U.S. call center to estimate the willingness to pay for alternative work arrangements relative to traditional office positions. A significant number of workers state that they are willing to give up a substantial share of their wages to avoid a schedule set by an employer on short notice. By comparison, the stated willingness to pay for choosing when to work is relatively low.

[33] provide complementary evidence from a revealed preferences approach to estimating workers' valuation of flexibility. They combine data from a natural field experiment conducted on a Chinese job board with survey and observational data. The experimental job ads differ randomly in offering jobs that are flexible regarding when and where one works. Both the survey evidence and the experimental estimates suggest that workers are willing to take lower pay for more flexible jobs. For instance, application rates are significantly higher for flexible jobs, conditional on the salary offered. [33] argue that a natural field experiment offering real jobs to real job seekers has several advantages over alternative approaches.[8] The participants in the natural field experiment are actually searching for jobs, properly incentivized to respond in ways most likely to get them the jobs they want, and unaware they were under scrutiny in a scientific study. The natural field experiment that we study

8. See [32] for a broader discussion of the advantages of natural field experiments.

have the same advantages. In addition, we can measure the participants' valuation of flexibility in terms of their actual work decisions, unlike the job board experiment that does not capture the final outcomes of the search process (such as callbacks for interviews, job offers, and actual remuneration).

Finally, methodologically, we join a set of recent studies that combines field experiments with structural methods to uncover key counterfactuals (see, e.g., [20]; [21]). In doing so, we highlight how the combination of theory and field experiments can be used to evaluate a wide range of economic issues (see also [42]).

2.2 Background and Experiment

We now review the labor market for Uber's drivers before describing the natural field experiment.

Uber Marketplace

Uber's rideshare platform is the largest service provider in the ride-sharing market in the U.S. In 2016, for example, it had a market share of about 83% of the U.S. consumer ride-sharing market. Uber and Lyft combined owned nearly 97 percent of this market. Uber connects riders and drivers through its app. Once a ride request is made, the app contacts nearby drivers for the ride. Drivers would see the rider's location. While drivers are incentivized to maintain a high acceptance rate, drivers can decide whether to accept this trip.

Drivers are effectively free to choose when and how much to work. There are no minimum-hours requirements and only modest constraints on maximum hours. Drivers are paid according to a fixed, non-negotiated formula. As described in detail later, workers earn a base fare per trip plus amounts for how long and how far they drive. On top of this standard fare, Uber offers fare multipliers when the demand for rides is sufficiently high compared to the supply of drivers (commonly referred to as surge pricing) or if the drivers are participating in the randomized experiment. Both the variation in surge multipliers and the randomness in the arrival of rider requests lead to variability in effective market wages per hour.

In most U.S. cities, there are relatively few barriers to becoming an Uber driver. While the exact requirements vary from city to city, drivers must typically fill out online paperwork, undergo a background check, and meet certain driver and vehicle requirements. In the years (2016-2018) and cities (Boston, Chicago, San Francisco) we consider, one of these requirements is that the vehicle's model year is 2001 or newer. By comparison, Lyft required that the vehicle must have a year model of 2003 or newer. As a robustness check, we take advantage of this difference in eligibility requirements to assess the importance of substitution between Lyft and Uber.

Experiment Setup: Guaranteed Surge Level

Our natural field experiments arises from the so-called Guaranteed Surge Levels (GSL hereafter). The GSL is essentially a fare multiplier that Uber randomly offers to a subset of drivers to increase their expected market wages during certain hours. Drivers who were active in the past 28 days and have completed at least 40 trips are eligible to receive fare multipliers through GSL.

The experimental setup is as follows. Uber divides each week into 2 blocks: Block 1 starts from Monday 4:00 a.m. and ends on Friday 3:59 a.m., while Block 2 goes from Friday 4:00 a.m. to 3:59 a.m. the next Monday. Over the course of a given block, GSL is switched on for a subset of hours. In Figure 2.1, we show an example of Block 1. During the example block, the highlighted hours, such as Monday 5:00–6:00 a.m., are chosen as experiment hours where drivers receive hour-specific fare multipliers. We refer to consecutive experiment hours as an experiment window. We refer to the schedule of GSL experiment hours within a block as a GSL menu. These menus vary across blocks. Drivers learn about the GSL menu, via email and/or the Uber app, the night before a block starts. Figure 2.2 shows an example of a GSL menu with multiple experiment windows, making clear to drivers when and for how long GSL will be switched on in an upcoming block.

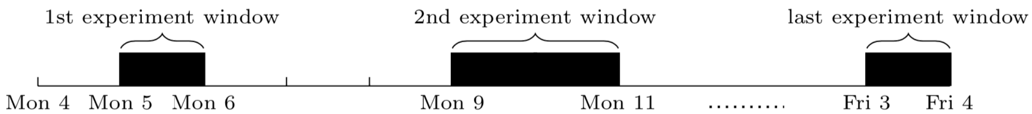


Figure 2.1: Example of a GSL Menu of Block 1 with Multiple Experiment Windows

Notes: Highlighted hours denote the hours when the GSL experiment is switched on.

For each block, eligible drivers are randomly assigned to treatment and control groups where the treatment group receives 0.1 higher GSL fare multipliers than the control group for all the experiment hours within the block. Consider again the example in Figure 2.1. Monday 5:00–6:00 a.m. and Monday 9:00–11:00 a.m. are two experiment windows. Suppose the control group drivers receive 1.1× fare multiplier in the first window and 1.3× fare multiplier in the second window. Since the treatment group always receives 0.1 higher GSL fare multipliers than the control group, the treated drivers would then be receiving 1.2× fare multiplier in the first window and 1.4× fare multiplier in the second window. At the end of each block, drivers are re-randomized into treatment and control groups for the next block.

On average, an experiment window lasts for about 5 hours, and there are about 7 experiment windows per block. Across blocks, there is variation in the days and hours of the experiment windows. In total, around 40 percent of the hours in our sample are subject to the GSL experiment. Thus, the GSLs generate considerable variation in expected wages at different days of the week and at various hours of the day.

Trip Earnings, Wages and Work Decisions

The unit of observation in our analysis is an individual driver at a given hour. Thus, we measure labor supply and wages on an hourly basis. We define total minutes worked per hour as the number of minutes for which a driver has the app on and is available for accepting

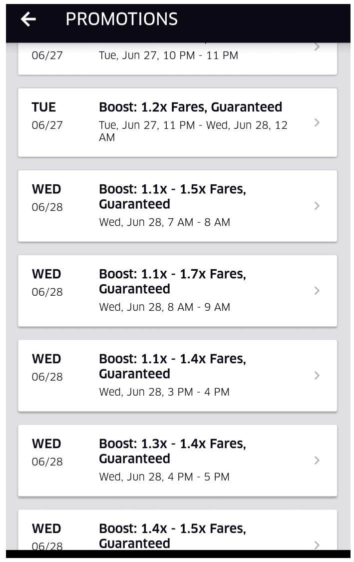


Figure 2.2: Example of a GSL Announcement to Drivers

Notes: The screenshot shows a GSL announcement to drivers on the Uber Driver App. The announcement clearly shows the hours when the GSL will be switched on and the corresponding GSL multiplier.

requests in that hour. In other words, a driver is said to be working if she is actively searching for rider requests.[9]

The observed hourly wage rate is measured as a driver's total earnings divided by the number of minutes worked in an hour, multiplied by 60. The total earnings in an hour are measured as the sum of trip earnings a driver receives, where each trip earnings is determined by the following formula:

$$\text{Trip Earning} = \max(\text{GSL}, \text{Surge}) \times \text{Baseline Fare}.$$

In this formula, GSL is the experimental fare multiplier, Surge is the demand-driven fare multiplier, and BaselineFare is the baseline trip earnings following Uber's fixed compensation

9. This is the same definition of working as in [14]. A driver is active if the driver-side app is turned on and she is available to accept requests for rides. This is to be distinct from a "browsing" mode in which the app is on but the driver has not indicated a willingness to accept rides.

rule:

$$\text{Baseline Fare} = \text{Fixed Price} + (\text{Price per Minute} \times \text{Minutes}) + (\text{Price per Mile} \times \text{Miles}).$$

As in [14], our measure of wages does not net out the variable costs of operating a vehicle. Therefore, our reservation wages should be interpreted as a gross quantity. It is important to observe, however, that labor supply decisions are based on the differences between expected wages and reservation wages, which do not depend on assumptions regarding the incorporation of time-invariant operating costs.

Drivers base labor supply decisions on expected hourly wages rather than the realized wages that we observe. To construct measures of expected wages, we predict the hourly wage a driver is likely to face in each hour. As a first step, we calculate the wage multiplier, defined as $\max(\text{GSL}, \text{Surge})$, from our detailed data on GSL and Surge. Next, we calculate the pre-multiplier wages as the observed hourly wages divided by the calculated wage multipliers. We then fit the following regression model to the panel data on pre-multiplier wage:

$$\tilde{W}_{it} = \alpha_i + \kappa_{h(t)} + \epsilon_{it}$$

where t is an hour, $h(t)$ is the hour of the week at t, i denotes a driver, α_i and $\kappa_{h(t)}$ are driver and hour-of-week fixed effects, and $\tilde{W}_{it}$ is the pre-multiplier hourly wage. For each hour t, we fit the model with the panel data up to $t-1$, and use the estimated $\hat{\alpha}_i$ and $\hat{\kappa}_{h(t)}$ to compute a predicted value for $\tilde{W}_{it}$ for every worker i in each hour t. The predicted hourly wage is then constructed as the product of the predicted pre-multiplier wage and the calculated wage multiplier.

To assess how well our prediction model performs, we compare it to alternative approaches using a cross-validation procedure with details in Appendix 2.A. This procedure repeatedly divides samples into a training sample and a testing sample. For each approach, we use the training samples to estimate a prediction model, and then we use the estimated model to form a prediction and calculate the out-of-sample mean squared errors on the testing samples. In addition to our current prediction model, we consider several alternative approaches, including a matching procedure with a K-means clustering method. Our prediction model performs considerably better than the alternative approaches.

2.3 Data and Experimental Findings

In this section, we describe the data and present the findings from the experiment.

Description of the Data

Our analyses are based on panel data of UberX and Uber Pool drivers who are eligible for the GSL experiment in Boston, Chicago, and San Francisco. In Boston and Chicago, we observe all these drivers. In San Francisco, we have data for a random subsample of 35 percent of the eligible drivers. In Boston and San Francisco, the duration of our data spans from October 2016 to March 2018. In Chicago, we have only one year of data, covering October 2016 to March 2017. For each driver, we observe gender, age, type of vehicle, minutes worked per hour, trip earnings, and fare multipliers.

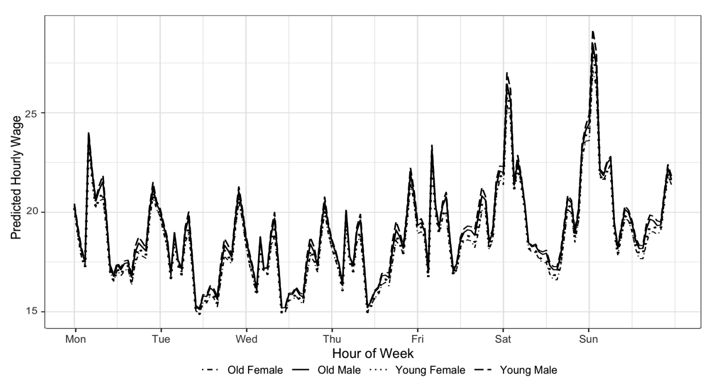


Figure 2.3: Predicted Hourly Wages Across Hours of the Week by Demographic Groups

Notes: We compute the average predicted wage at every hour of the week and for each demographic group.

Our sample covers 333,172 drivers. A quarter of these are female, and the median age is about 38. The average observed wage is $19.29 per hour. In Figure 2.3, we plot the predicted hourly wage over time according to gender and age. We define drivers as young if they are younger than 38 years old. Consistent with [14], most of the heterogeneity in hourly wages is due to hours of the day. It is also evident that hourly wages tend to be higher in the weekends and that male drivers have only slightly higher hourly wages as compared to female drivers.

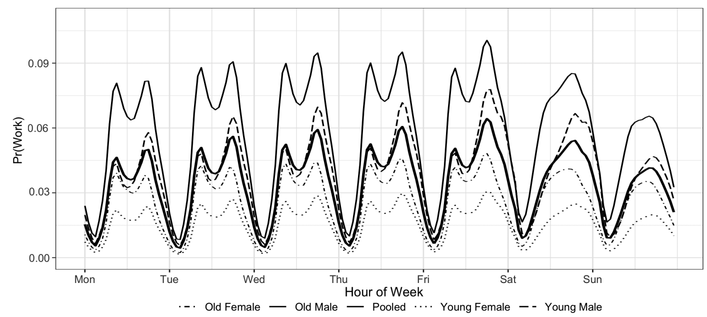


Figure 2.4: Probability of Working Across Hours of the Week by Demographic Groups

Notes: In this figure, we compute the share of active drivers at every hour of the week and for each demographic group. We define active drivers as those who work any positive number of minutes in a given hour. "Pooled" refers to the combined sample across demographic groups.

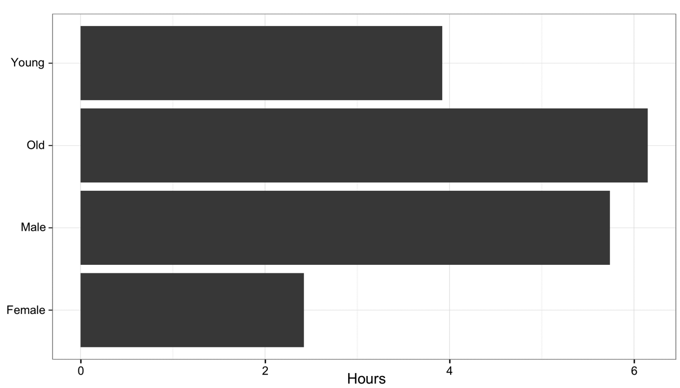


Figure 2.5: Average Number of Hours Worked per Week by Demographic Groups

Notes: In this figure, we compute the average number of hours worked for each demographic group. An hour worked is defined as an hour where a driver works for any positive number of minutes.

Figure 2.4 plots the probability of working over time for all drivers and by subgroup. The work probability varies considerably across hours of the week and days of the week. There are also distinct differences in the probability of working by age and gender. Conditional on age, male drivers are much more likely to work, especially during the daytime. Holding gender fixed, old drivers tend to work more than young drivers. In Figure 2.5, we show the average number of hours worked per week by age and gender. In expectation, drivers work nearly five hours per week. However, males drivers work twice as many hours per week as female drivers, and young drivers work 36 percent less than older drivers.

Checking Covariate Balance

In a properly implemented, randomized experiment with a sufficiently large sample size, we expect the treatment and control groups to be balanced in their distribution of pre-treatment variables. To assess this, we check the covariate balance by regressing the treatment status in the GSL experiment on the pre-treatment characteristics of the drivers:

$$D_{it} = X_{it}'\beta + u_{it}$$

where D_{it} is an indicator variable that is equal to 1 if driver i is assigned to the treatment group in the GSL experiment at time t, and X_{it} is a vector of covariates that include gender, age, number of trips completed, and past wages. All these covariates are measured in the week before the randomization. Table 2.1 reports the estimates. In Column 1, we regress treatment status on each characteristic separately. There is no evidence of systematic differences between drivers in the treatment and control groups in the characteristics considered. In Column 2, we regress treatment status on all the characteristics in a multiple regression. Consistent with the randomization, we cannot reject the null hypothesis that all coefficients are zero.

Pre-determined Var.	(1) **Separate Regression** Var. on Treatment	(2) **Joint Regression** Treatment on Var.
Female (×10)	-0.0039	0.0000
	(0.0025)	(0.0001)
Age (×100)	-0.0003	-0.0651
	(0.0009)	(0.0426)
Wage Last Week (×100)	0.0051	-0.0000
	(0.0025)	(0.0000)
Trips Completed (×1000)	-0.0000	0.0536
	(0.0000)	(0.0259)
N (Blocks)	22,318,255	22,318,255
R^2		0.0000
F Statistic		1.8450
p value - F		0.1171

Table 2.1: Balance Tests

Notes: In Column 1, we regress each driver characteristic on the treatment status separately, and each row represents a separate regression. In Column 2, we regress the treatment status on all the characteristics in one regression. We use the F-test to examine whether one or more of the coefficients of these four pre-determined variables are significantly different from zero. For interpretability, we scale the regression coefficients and the standard errors of the driver's gender by 10, the driver's age and past wages by 100, and the driver's total number of trips completed by 1000.

First Stage: Effects of GSL on Expected Market Wages

Figure 2.6 presents the treatment effects of the GSL experiment on the expected market wages during the experiment hours. These treatment effects are obtained by OLS estimation of the predicted hourly wage on a dummy variable of being in the treatment group. These estimates form the first stage in the IV estimation of the effects on labor supply of exogenous changes in predicted wages. The shaded area in the figure indicates the hours with the experiment switched on. The figure shows that assignment to the GSL experiment increases the average predicted hourly wage by around forty cents or, equivalently, around 1.8 percent.To examine if the wage effects are persistent across experiment hours, we divide each experiment window in half, and then estimate the treatment effects separably for each half. The estimates suggest little, if any, changes in the wage effects across hours within the experiment window.

Reduced Form: Labor Supply Responses to GSL

We now document that drivers respond to the GSL experiment both during the experiment windows and in the hours preceding and following these periods. This is done by OLS estimation of labor supply on a dummy variable of being in the treatment group. These effects will be the reduced form estimates in the IV estimation of the effects on labor supply of exogenous changes in predicted wages. We consider labor supply responses along two margins: In each hour, we measure if the driver worked at all and the number of minutes she worked.

In Figure 2.7, we plot the effects of the GSL experiment on the labor supply responses of the drivers before, during, and after the experiment window. These effects are represented by the solid lines. The dotted lines represent the changes in wages due to the GSL experiment. We find significant changes in labor supply during the experiment window. During this window, the treated drivers increase the employment rate per hour and the hours of work by about one percent as compared to the control drivers. There is also some suggestive evidence of persistent effects outside the experiment window. For example, in the hour following the experiment window, the labor supply is half a percent higher for the treated group as compared to the control group.

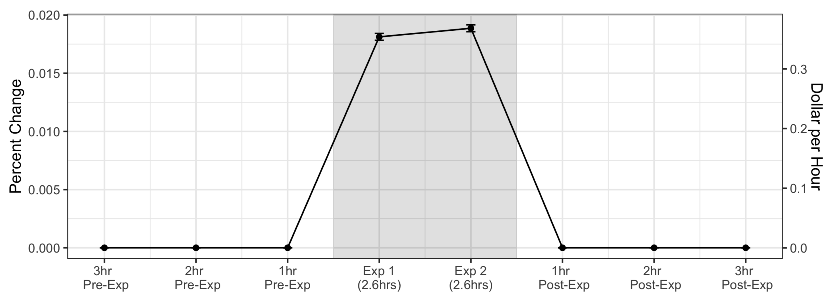
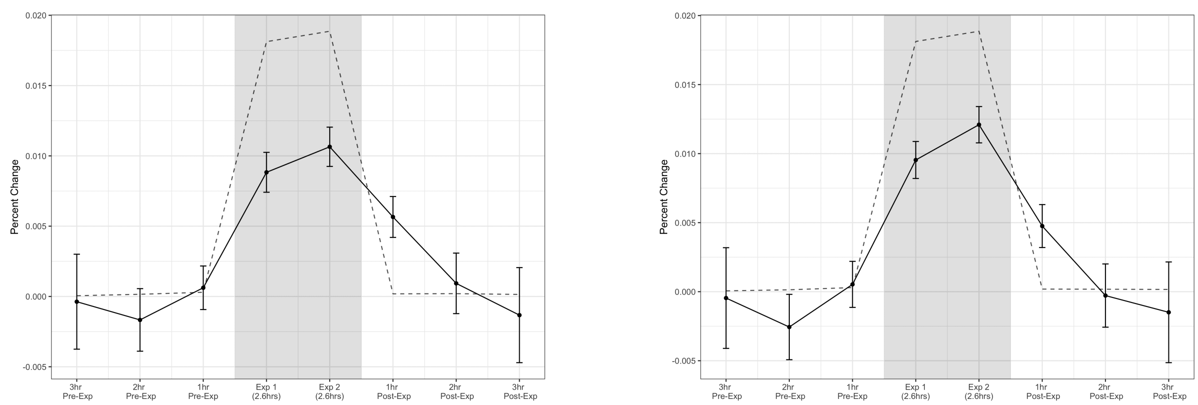


Figure 2.6: Treatment Effects of the GSL Experiment on the Expected Market Wages

Notes: In this figure, we present the estimates of the GSL experiment on the expected market wages during the experiment hours (in the hours before and after the experiment window, the effect is zero). The estimation is performed separately for the first half (Exp 1) and the second half (Exp 2) of the experiment windows. On average, the experiment windows last 5.2 hours. The shaded area in the figure indicates the experiment hours with the GSL switched on. The y-axis on the left shows percent differences between the treatment and the control groups. It is computed as the treatment effect on the predicted hourly wage divided by the predicted hourly wage of the control group. The y-axis on the right reports the estimated effects in dollars per hour. The bars indicate the 90% confidence intervals calculated from subsampling bootstrap.



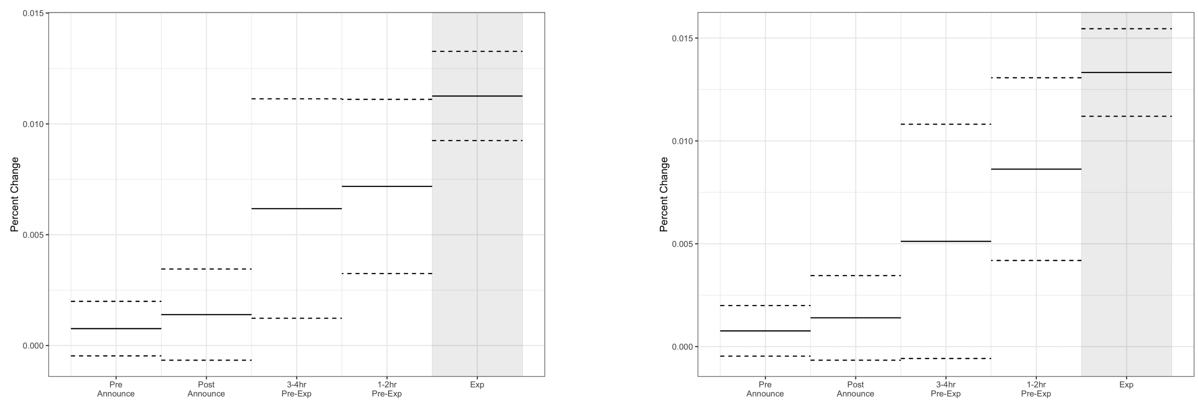
(a) Working

(b) Minutes Worked

Figure 2.7: Treatment Effects on Labor Supply Responses

Notes: In this figure, we illustrate the changes in the probability of working and minutes worked. We recenter all the experiment windows and plot the x-axis the same way as in Figure 2.6. The estimation is performed separately for the first half (Exp 1) and the second half (Exp 2) of the experiment windows. On average, the experiment windows last 5.2 hours. The shaded area in the figure indicates the experiment hours with the GSL switched on. The solid line represents the difference in labor supply responses, and the dashed line represents the difference in predicted hourly wages. The bars indicate the 90% confidence intervals.

When interpreting the estimated effects in the hours preceding and following the experiment window, it is important to recognize that most blocks have several GSL experiment windows. As a result, the hours preceding and following a given experiment window are likely to be confounded by other GSL experiment windows. For example, in 30 percent of the blocks in our sample, there exists at least two experiment windows that are no more than an hour apart. To address this concern, we take advantage of the re-randomization at the start of every block and restrict attention to the experiment windows preceding and following the re-randomization. By estimating the labor supply responses around these experiment windows, we avoid confounding anticipatory and persistent responses with other GSL experiment windows.

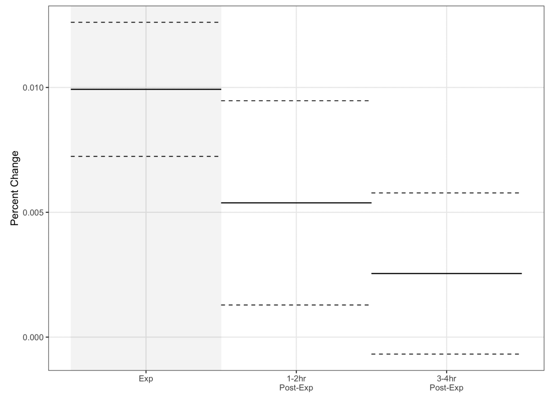


(a) Working

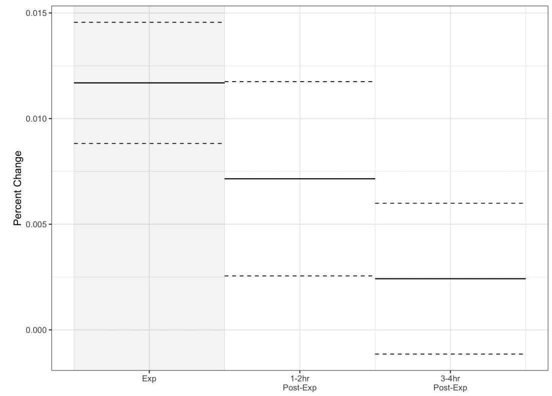
(b) Minutes Worked

Figure 2.8: Treatment Effects on Labor Supply Responses in First Experiment Windows

Notes: We illustrate the labor supply responses immediately after re-randomization. We pool all the first experiment windows for estimation. The point estimates are constructed in the same way as Figure 2.7. The shaded area in the figure indicates the experiment hours in the first experiment windows. The "Pre Announce" period contains all hours before the announcement of the GSLs, while the "Post Announce" period contains the hours after announcement up to 5 hours before the first experiment windows. The dashed bars indicate the 90% confidence intervals.



(a) Working



(b) Minutes Worked

Figure 2.9: Treatment Effects on Labor Supply Responses in Last Experiment Windows

Notes: We illustrate the labor supply responses immediately before re-randomization. We pool all the last experiment windows for estimation. The point estimates are constructed in the same way as Figure 2.7. The shaded area in the figure indicates the experiment hours in the last experiment windows. The dashed bars are the 90% confidence intervals.

In Figures 2.8 and 2.9, we present the estimates from the experiment windows preceding and following the re-randomization. Even four hours prior to the experiment window, we see the treated drivers are more likely to be working as compared to drivers in the control group. The anticipatory response is most pronounced in the hour just before the experiment window. The same holds true for the persistent responses. In the first few hours following the experiment, the labor supply of the treated drivers remains significantly higher than the drivers in the control group. As time passes, these differences decline, and four hours after the experiment window the treated drivers work as much as the control drivers.

IV Results: Labor Supply Responses to Exogenous Changes in Expected Market Wages

We now turn attention to how the labor supply of drivers responds to exogenous changes in expected market wages. This is done by 2SLS regression with labor supply as the dependent variable, predicted wages as the treatment variable, and the experiment as the instrument.

The resulting 2SLS estimates correspond to the ratio of the reduced form and the first stage estimates reported above. Table 2.2 presents the IV estimates. In the first three columns, we measure labor supply as any work in a given hour. The last three columns measure labor supply as minutes worked during an hour. In Columns 1 and 4, we use the data from all the experiment hours. Columns 2 and 5 restrict attention to the hours surrounding the first experiment window in a block, whereas Columns 3 and 6 consider only the hours surrounding the last experiment window in a block.

The first stage estimates in Columns 1 and 4 are very precise and show that the GSL experiment raises the predicted hourly wage by nearly forty cents. More importantly, the IV estimates in these columns imply that a $10 increase in hourly wages would raise the share of workers that drive in a given hour by 1.4 percentage points or, equivalently, by 27 percent. By comparison, this increase in hourly wages would raise the amount of minutes that an average driver works by about 30 percent, from 2.26 to nearly 3 minutes per hour. The columns other than 1 and 4 quantify the labor supply responses of the drivers in the hours preceding and following the exogenous changes in predicted wages. We find that these anticipatory and persistent responses are significant and economically relevant. For both measures of labor supply, the anticipatory and persistent responses are about a third of the size of the responses during the experiment hours.

To better understand what drives the increase in labor supply, it is useful to decompose the IV estimates into responses on the extensive (any work in a given hour) and the intensive margin (minuted worked conditional on working in a given hour). Concretely, we decompose the estimate in Column 4 of Table 2.2 as follows:

$$\underbrace{\frac{\partial}{\partial W_t} E(\text{MinutesWorked}_t)}_{0.680} = \underbrace{\frac{\partial}{\partial W_t} E(\text{MinutesWorked}_t | \text{Work}_t) \times Pr(\text{Work}_t)}_{0.064\ (9.4\%)}$$
$$+ \underbrace{E(\text{MinutesWorked}_t | \text{Work}_t) \times \frac{\partial}{\partial W_t} Pr(\text{Work}_t)}_{0.616\ (90.6\%)}$$

where Work_t is an indicator variable that is equal to one if the driver works in a given hour. Using the drivers in the control group, we calculate $E(\text{MinutesWorked}_t|\text{Work}_t)$ and $Pr(\text{Work}_t)$ while $\frac{\partial}{\partial W_t} E(\text{MinutesWorked}_t)$ and $\frac{\partial}{\partial W_t} Pr(\text{Work}_t)$ are taken from the estimates in Columns 4 and 1, respectively. The results suggest that responses at the extensive margin account for nearly all the increase in the amount of minutes worked during an hour. This finding suggests it is important to model the driver's decision to work or not in a given hour, not the amount of minutes she works within an hour.

Heterogeneity in Labor Supply Responses

So far, we have focused on the average labor supply responses across all drivers. However, these average impacts miss a lot: The labor supply responses vary systematically both across people and over time. In order to discover these heterogeneous effects, we apply the method of [15] to estimate an IV model with a full set of interactions between the endogenous regressor of interest, wages, and the pre-determined covariates, gender, hour of the day, and day of the week. Since there are 168 hours per week, we have 168 bins for the time dimension. For age, we divide the sample into four equally sized groups: Younger than 30 years, 30 to 38 years, 38 to 48 years, and older than 48 years. Since gender is binary, we therefore get 1,344 mutually exclusive and collectively exhaustive groups.

We sort these 1,344 IV estimates in an increasing order. In Figure 2.10 we illustrate these sorted IV estimates. The x-axis represents the percentile rank in the distribution of

Dependent Var.	Working			Minutes Worked		
Panel A: IV						
	All	First	Last	All	First	Last
Experiment Hrs	0.0142	0.0167	0.0113	0.6801	0.8130	0.5558
	(0.0014)	(0.0009)	(0.0009)	(0.0654)	(0.0382)	(0.0394)
Anticipation		0.0053			0.2451	
		(0.0009)			(0.0447)	
Persistence			0.0036			0.1800
			(0.0009)			(0.0403)
Control Mean	0.0525	0.0474	0.0375	2.2600	1.9484	1.5636
of Dep. Var.	(0.0002)	(0.0002)	(0.0002)	(0.0084)	(0.0088)	(0.0075)
Panel B: First Stage						
	All	First	Last	All	First	Last
First Stage	0.3608	0.3193	0.3288	0.3608	0.3193	0.3288
	(0.0037)	(0.0057)	(0.0062)	(0.0035)	(0.0054)	(0.0061)
Control Mean	19.5091	20.3611	19.7459	19.5091	20.3611	19.7459
of First Stage	(0.0138)	(0.0196)	(0.0189)	(0.0133)	(0.0186)	(0.0180)

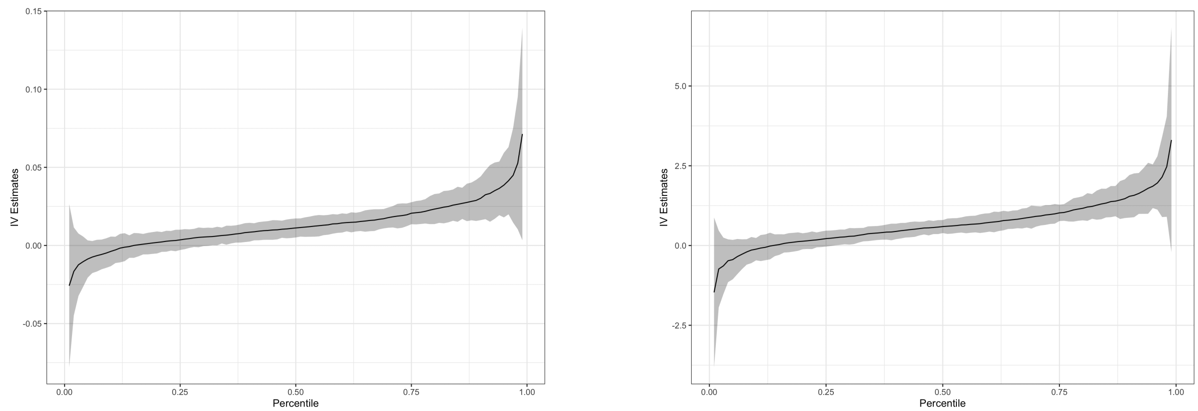
Table 2.2: Labor Supply Responses: IV and First Stage Estimates

Notes: "Control mean" is the expected outcome for the control group. IV is estimated as the increase in the probability of working per hour (or minutes worked per hour) per $10 increase in the predicted hourly wage. The standard errors of the IV estimates and the first stage are estimated by bootstrap. Column 2 and Column 5 show the estimates for the first experiment windows after re-randomization. Column 3 and Column 6 show the estimates for the last experiment windows before re-randomization.

the estimated effects, and the y-axis shows the estimated effect sizes. Panel (a) graphs the estimated effects for working and Panel (b) shows the estimated effects for minutes worked. It is evident that the IV estimates vary widely across groups. At the 25th percentile, the estimated effect for working is 0.0036, while the value at the 75th percentile is 0.021, almost 6 times as large. The estimated effect for minutes worked at the 75th percentile is almost 5 times as large as the value at the 25th percentile.

A natural question is what drives the heterogeneity in labor supply responses documented in Figure 2.10. To answer this question, we begin by examining the time dimension. In Figure 2.11a, we plot the average responses across all drivers for different hours of the day and for weekends versus weekdays. There is substantial heterogeneity along the time dimension.

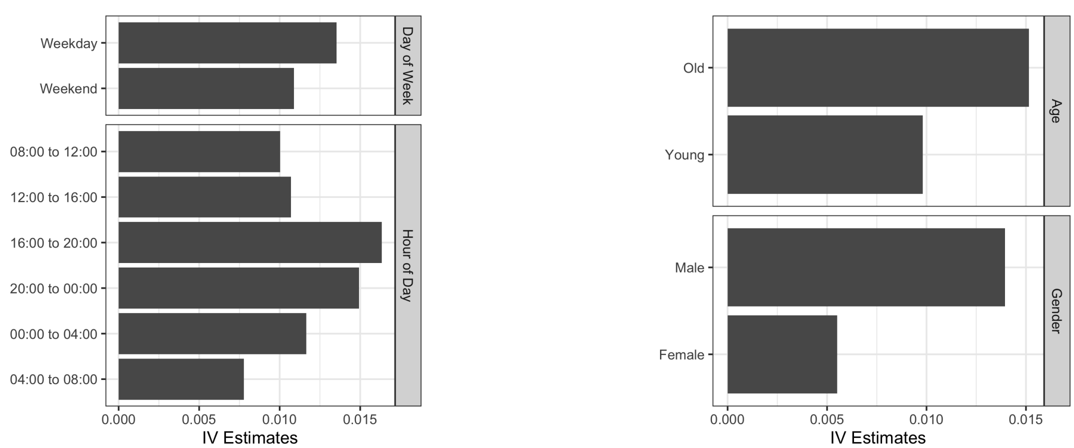
Most of this variation comes from the differences across hours of the day rather than days of the week. When performing formal statistical tests, we can strongly reject the null hypotheses of equal average responses across hours of the day and across weekends versus weekdays.



(a) Working (b) Minutes Worked

Figure 2.10: Heterogeneity in IV Estimates

Notes: In this figure, we plot the distribution of the heterogeneous treatment effects. The solid line indicates the IV estimates measured as the increase in the probability of working (or minutes worked) w.r.t. a $10 increase in the predicted hourly wage. The estimates and the 90% bootstrap uniform confidence bands are derived following [15] based on the linear model with full saturation of observed heterogeneity in hours of the week, young versus old, and male versus female.



(a) Over Time (b) Across Demographic Groups

Figure 2.11: Heterogeneity in IV Estimates over Time and Across Demographic Groups

Notes: Figure (a) compares the IV estimates across hours of the day, and weekdays versus weekend. Figure (b) compares the IV estimates across demographic groups. The unit of the IV estimates is the increase in the probability of working w.r.t. a $10 increase in the predicted hourly wage.

As shown in Figure 2.11b, there is also considerable heterogeneity by gender and age. On average, male and old drivers have larger responses than young and female drivers. The differences across gender are larger than those by age. Again, the null hypotheses of equal average responses by gender or age are strongly rejected in the data.

We conclude the analysis of heterogeneity by examining how much of the heterogeneity in the estimated labor supply responses can be explained or accounted for by various covariates. We begin by regressing the 1,344 IV estimates reported in Figure 2.10 on indicator variables for time. In Table 2.3, We find that day of the week explains as little as 1.2 percent of the variation in the labor supply responses. By comparison, measuring time through indicators for hours of the day increases the R-squared to 12.4 percent. There are only small gains in explanatory power from including indicators for both hour of the day and day of the week in a separable fashion. By way of comparison, interactions between hours of the day and days of the week are empirically important to explain the pattern of heterogeneity in labor supply responses, increasing the R-squared from 13.7 percent to 34.9 percent. Also including indicator variables for gender and young (defined as younger than 38) further increases the R-squared by a few percentage points. By comparison, a flexible regression model with interactions between the day of the week, the hour of the day, gender and young explains nearly 70 percent of the variation in labor supply responses. Taken together, these results suggest it is key to let preferences of the drivers vary by gender and age, and for each demographic group, across hours of the day and days of the week.

Independent Variable	Explanatory Power (R^2) Working	Minutes Worked
Day of Week	0.012	0.012
Hour of Day	0.124	0.124
Day of Week + Hour of Week	0.137	0.137
Day of Week × Hour of Day	0.349	0.349
(Hour of Day × Day of Week) + (Young × Gender)	0.384	0.384
(Hour of Day × Day of Week) × (Young × Gender)	0.691	0.718

Table 2.3: Explanatory Power in Regressions of Labor Supply Responses on Covariates for Time and Demographics

Notes: We regress the IV estimates of labor supply responses on the time, age, and gender dummies, and we report the R-squared in this table. Each number in this table corresponds to a separate regression. Each regression is weighted by the inverse of the variance of the IV estimates.

Comparison with OLS estimates

Table 2.4 compares OLS estimates of the labor supply responses to the IV estimates we obtain using the experiment. These results suggest that unobserved determinants of wages, if ignored, lead to a significant downward bias in the estimated labor supply responses. In particular, the OLS estimates in Column 1 show much weaker associations between labor supply and wages than the IV estimates. This downward bias is consistent with demand being high when it is costly for the drivers to work. Including fixed effects for workers, days of the week, and hours of the day reduces the bias, as shown in Column 2. However, the labor supply elasticities remain too small. This finding suggests that idiosyncratic factors, such as weather conditions and entertainment events, may create high demand while, at the same time, make driving more costly or difficult.

Dependent Variable: Working			
	OLS	FE	IV
Estimates	0.0099	0.0125	0.0142
	(0.0003)	(0.0003)	(0.0014)
	OLS = IV	FE = IV	
p-value	0.0001	0.1143	
Dependent Variable: Minutes Worked			
	OLS	FE	IV
Estimates	0.5223	0.6296	0.6801
	(0.0157)	(0.0180)	(0.0654)
	OLS = IV	FE = IV	
p-value	0.0005	0.2066	

Table 2.4: OLS and IV Estimates of Labor Supply Responses

Notes: In this table, we report the OLS, FE, and IV estimates of the labor supply responses. OLS is estimated by regressing the outcome variables on predicted hourly wages. The standard errors for the OLS estimates are clustered at the drivers level, and the standard errors for the IV estimates and the p-values are estimated by bootstrap. The fixed effects include the driver fixed effect and the hour of week fixed effect.

2.4 Dynamic Model of Labor Supply

The experimental estimates provide key data points for learning about labor supply elasticities, reservation wages and the value of job flexibility, but do not by themselves tell us these quantities. In order to recover the labor supply elasticities and reservation wages and to infer the value of job flexibility, we now develop, identify, and estimate a dynamic model of labor supply. In this section, we present this model and discuss the parameter estimates. In Appendix 2.C, we compare these estimates to those produced by more restrictive models, including a static labor supply model. This comparison highlights how several of our modeling choices – including adjustment costs, permanent observed and unobserved heterogeneity, and the field experiment to address wage endogeneity – are key not only to match the data but also for the estimates of the reservation wages and for the results from the counterfactual analyses.

Model Setup

Driver's Problem

We model the driver as living infinitely many periods where each period is an hour. In each period t, the driver decides whether to work $a_{it} = 1$ or rest $a_{it} = 0$, taking into account both the current period payoff $U_{it}(a_{it})$ and how her choice in t will affect the payoffs in the future $\tau > t$. In each period t, a driver chooses a_{it} in order to maximize the expected sum of discounted flow payoffs:

$$\max_{a_{it}} \mathbb{E}\left[\sum_{\tau=t}^{\infty} \rho^{\tau} U_{i\tau}(a_{i\tau})|a_{it}\right]$$

where i indexes a driver, τ indexes hour, $U_{i\tau}(a_{i\tau})$ is the flow payoff associated with choice $a_{i\tau}$, and ρ is the discount rate. The expectation is taken over the future values of $U_{i\tau}(a_{i\tau})$ given the current choice a_{it} for $\tau \geq t+1$.

4.1.2 Preferences

We sort drivers into subgroups based on their age and gender, $X = (\mathbf{1}\{\text{Female}\}, 1\{\text{Young}\})$. For a driver in an observed subgroup $X = x$ who works in a given city, the flow payoff associated with action a_{it} is given by

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \eta_{j(i),h(t)} + \xi_{it} + \epsilon_{1it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

where i is a driver, t is a calendar hour (e.g., 2018/10/10, 9 a.m.) and $h(t)$ is an hour of a week at time t (e.g., Monday 9 a.m.). We also include city fixed effect in the empirical specification of U_{it} to allow for systematic differences in the costs of driving across cities. For notational simplicity, we suppress these fixed effects as well as the conditioning on X.

A driver's flow payoff from work depends on the wage she may earn, w_{it}, and the time-specific shifter of the cost of driving at a given hour of the week, $\beta_{h(t)}$. The empirical counterpart of w_{it} is the predicted hourly wage as described in Section 2.2. If the driver did not work at $t-1$, she needs to pay an adjustment cost to start to work, μ. The parameter $\eta_{j(i),h(t)}$ captures the unobserved type j of driver i at time $h(t)$. For example, full-time drivers who are more likely to drive at all times have higher $\eta_{j(i),h(t)}$ at all $h(t)$ than infrequent drivers, while evening drivers have higher $\eta_{j(i),h(t)}$ only in the evenings. Driver type $\eta_{j(i),h(t)}$ is known to the drivers themselves but unobserved to the analyst.

A driver's flow payoffs from the choices at t are also affected by a set of choice-specific preference shocks, ξ_{it}, ϵ_{1it} and ϵ_{0it}, which are revealed to the driver at the beginning of t. The component ξ_{it} captures the unobserved preference shocks that may correlate with wages. For example, an individual may dislike to drive in periods with heavy traffic or poor weather and these conditions may also covary with the demand for rides and thus offered wages. The components ϵ_{1it} and ϵ_{0it} capture the idiosyncratic preference shocks that are independent of wages.

The equation for the predicted hourly wage of worker i in t is specified as follows:

$$w_{it} = \delta_{h(t)} + \delta_0 m_t + \delta_1 m_t \times z_{it} + u_{it}$$
$$u_{it} \sim G_u(u_{it}|u_{it-1}, a_{it-1}, z_{it-1}, h(t-1), m_{t-1}, c(i))$$

where m_t is an indicator of whether t is an experiment hour and z_{it} is an indicator for being assigned to the treatment group in an experiment hour, $\delta_{h(t)}$ is a fixed effect for hours of week, and u_{it} represents the unobservable determinants of wages. By including m_t in the wage equation, we allow the wage to be different when a GSL experiment is switched on. The parameter δ_1 captures the exogenous change in the wage for the treated drivers during experiment hours.

The wage in our model evolves as a first-order Markov process. We allow persistence in wages by letting u_{it} depend on u_{it-1}, together with lagged choice a_{it-1}, lagged treatment z_{it-1}, lagged experiment hour m_{t-1} and city $c(i)$. We allow a driver's unobserved type $\eta_{j(i),h(t)}$ to depend nonparametrically on her initial wage and work decision. The endogeneity of wages arises if $Cov(\xi_{it}, u_{it}) \neq 0$. Thus, an exogenous wage process is a special case of our model in which the costs of working do not covary with market wages, $\delta_0 = \delta_1 = 0$.

4.1.3 Timeline and Information Set

Recall that at 4 a.m. every Monday and Friday, drivers are randomized into treatment and control groups. Drivers are then informed of when and for how long GSL will be switched on in an upcoming block. We specify the timeline within and between blocks as follows.

Timeline Between Blocks. At the beginning of each block, a driver i learns her treatment status z_{is} and the experiment hours m_s for all hours s in the block. At the same time, the driver forms expectations about her treatment status and the experiment hours in future blocks. The driver then sequentially makes labor supply decisions, taking into account the current flow payoff and the continuation values. The drivers are re-randomized at the start of the next block. Figure 2.12 presents an example of the timeline between blocks.

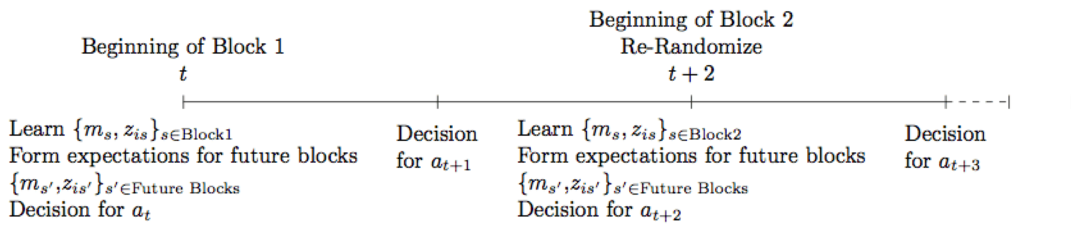


Figure 2.12: Example of Timeline with 2 Blocks and 2 Periods per Block

Timeline Within a Given Block. The decision timeline within a block is given as follows. Let t be an arbitrary period within a given block. At the beginning of t, driver i learns the realization of u_{it} (equivalently, her market wage w_{it}) and the realizations of the preference shocks $(\xi_{it}, \epsilon_{1it}, \epsilon_{0it})$. Based on these realizations, she forms expectations about future values of $u_{it'}, \xi_{it'}, \epsilon_{1it'}, \epsilon_{0it'}$. Next, she makes the work decision for period t, taking into account the current flow payoffs and how her decision at t will affect her future flow payoffs.

Assumptions and Identification

Our identification argument combines a control function based on the experiment with fairly standard assumptions in the dynamic discrete choice literature. In this section, we briefly discuss the key assumptions and the outline of the identification argument. The details are in Appendix 2.B.

The identification argument begins by making the following assumptions:

Assumption 1. *Control Function Assumption*

(Instrument Exogeneity) $z_{it} \perp (\epsilon_{0it}, \epsilon_{1it}, \xi_{it}, u_{it})$

(Joint Normality) $\begin{pmatrix} u_{it} \\ \xi_{it} \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho_{u\xi} \\ \rho_{u\xi} & 1 \end{pmatrix})$

Under joint normality of u_{it} and ξ_{it}, we can rewrite ξ_{it} as

$$\xi_{it} = \frac{\rho_{u\xi}}{\sigma} u_{it} + \psi_{it} = \frac{\rho_{u\xi}}{\sigma}(w_{it} - \delta_{h(t)} - \delta_0 m_t - \delta_1 m_t \times z_{it}) + \psi_{it}$$

where $\psi_{it} \sim \mathcal{N}(0, 1 - \rho_{u\xi}^2)$ and $\psi_{it} \perp u_{it}$ by construction. We define a new state variable $\phi_{it} \equiv w_{it} - \delta_{h(t)} - \delta_0 m_t - \delta_1 m_t \times z_{it}$. Thus, the flow payoffs of the problem become

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \eta_{j(i),h(t)} + \frac{\rho_{u\xi}}{\sigma}\phi_{it} + \nu_{it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

$$\begin{aligned} w_{it} &= \delta_{h(t)} + \delta_0 m_t + \delta_1 m_t \times z_{it} + u_{it} \\ u_{it} &\sim G_u(u_{it} | u_{it-1}, a_{it-1}, z_{it-1}, h(t-1), m_{t-1}, c(i)) \end{aligned}$$

where $\nu_{it} = \psi_{it} + \epsilon_{1it}$.

In addition to Assumption 1, we make the following set of assumptions which are often invoked in the literature on dynamic discrete choice (see [54], [45], [37]):

Assumption 2. *Standard dynamic discrete choice assumptions*

(IID) $\epsilon_{0it}, \epsilon_{1it}$ are iid across i, t

(EXOG) $\nu_{it} \perp (w_{it}, \eta_{j(i),h(t)}, \phi_{it})$

(CI-X) State transition probability F satisfies

$$F(w_{it+1}, \phi_{it+1} | a_{it}, a_{it-1}, w_{it}, \phi_{it}, h(t), j, \nu_{it}) = F(w_{it+1}, \phi_{it+1} | a_{it}, a_{it-1}, w_{it}, \phi_{it}, h(t))$$

(DISTR) Distributional assumption on ϵ_{0it} and ν_{it}, and independence: $\epsilon_{0it} \perp \nu_{it}$.

(DISCOUNT) ρ is known

(REACH) All states are reachable at any given time

$$F(w_{t+1}, \phi_{t+1} | a_t, a_{t-1}, w_t, \phi_t, h(t), j) > 0 \;\forall a_t, a_{t-1}, w_t, \phi_t, h(t), j$$

(NTYPE) The number of types is small.

Assumption 2 follows the standard assumptions in the dynamic discrete choice literature. The only difference is that, in our setup, the state-dependent unobserved preference shock, ξ_{it}, can be re-written as a combination of an observed state, ϕ_{it}, and an idiosyncratic component ψ_{it}, as a result of Assumption 1. Under the restrictions (IID), (EXOG), (CI-X), (DISTR), (DISCOUNT), we can identify the parameters in the flow payoff in the absence of unobserved heterogeneity (see [54], [45]). The restrictions (REACH) and (NTYPE) are made to incorporate unobserved heterogeneity in the model. Under these restrictions, the structural parameters in the flow payoffs are identified (see [34], [37]). The restriction (REACH) imposes that the entire support of wages at $t+1$ has a positive probability conditional on the state at t. In our problem, the restriction requires that at each hour of the week and conditional on the lagged choice and the wage in the previous period, a driver may get any wage in the support with a strictly positive probability. The restriction (NTYPE) limits the number of unobserved types among drivers. We allow for three unobserved types of drivers. For the restriction (DISTR), we assume ϵ_{0it} and ν_{it} are both distributed as T1EV.

It is useful to observe that some of the restrictions in Assumption 2 are not that strong in our setting. For example, the restriction (CI-X) implies that, conditional on choices, the transition probability of the state variables is the same across unobserved types and independent of transitory shocks. To understand this restriction, consider a driver who gets tired when working in period t. As a result, she may drive slower or take fewer trips and thus earn less if she chooses to continue driving in t+1. Restriction (CI-X) permits such a scenario. The restriction (DISCOUNT) requires the discount rate to be known to the analyst. In our setting, the data suggests that temporal decisions are primarily driven by heterogeneity in preferences over when to work $\beta_{h(t)}$, not discounting of future payoffs over a relatively short period of time. Thus, we think the discount rate plays a minor role for the behavior we observe. Assuming an annual interest rate as 5 percent, we set the hourly discount rate $\rho = 1/(1 + \frac{0.05}{365 \times 24})$.

4.2.1 Value Function and Structural Equation

We now describe the solution to the model and the structural equation. Under the standard conditions, the driver's problem can be characterized by the Bellman equation:

$$V(s_{it}) = \max\{V(a_{it} = 1, s_{it}), V(a_{it} = 0, s_{it})\} \quad (2.1)$$

where s_{it} is a vector of observed state variables and the unobserved type, $V(s_{it})$ is the (ex-ante) value function for driver i who is in state s_{it}, and the choice-specific conditional value functions $V(a_{it} = 1, s_{it})$ and $V(a_{it} = 0, s_{it})$ are defined as follows:

$$V(a_{it} = 1, s_{it}) = s_{it}'\theta + \xi_{it} + \epsilon_{1it} + \rho \mathbb{E} V(s_{it+1})$$

$$V(a_{it} = 0, s_{it}) = \epsilon_{0it} + \rho \mathbb{E} V(s_{it+1})$$

where $\theta = (\gamma, \beta_{h(t)}, \mu, \eta_{j(i),h(t)})$ is a vector of structural parameters in the flow payoff of work. The expectation is taken over future states and actions, $s_{i\tau}$ and $a_{i\tau}$ $\forall \tau \geq t+1$, as well as future preference shocks $(\epsilon_{0i\tau}, \epsilon_{1i\tau}, \xi_{i\tau})$ $\forall \tau \geq t+1$, conditional on s_{it} and a_{it}, according to the evolution of states $F(s_{it+1}|s_{it}, a_{it})$ and the distribution of shocks. Thus, with equation (2.1), we can describe the driver's decision rule as follows: At the beginning of each period t, driver i learns state s_{it}, and chooses to work if and only if $V(a_{it} = 1, s_{it}) \geq V(a_{it} = 0, s_{it})$.

4.2.2 Estimation

The first step in our estimation procedure is to use OLS to estimate the parameters of the wage equation, $\delta_{h(t)}$, δ_0, δ_1, and then to obtain the empirical counterpart of u_{it} as measured by the residuals. We then non-parametrically estimate the transition probability of u_{it} given lagged state variables $(u_{it-1}, a_{it-1}, z_{it-1}, h(t-1), m_{t-1})$.

We follow the two-stage estimator of Arcidiacono and Miller (2011) to estimate the model with unobserved types by maximizing the log likelihood of the finite mixture model:

$$\{\hat{\theta}, \hat{\pi}\} = arg\, max_{\theta,\pi} \sum_{i=1}^{N} \ln[\sum_{j=1}^{J} \pi_j(s_{i1}) \prod_{t=1}^{T_i} l(a_{it}|s_{it}, j, \hat{p}, \theta)] \tag{2.2}$$

where $s_{it} = (w_{it}, \phi_{it}, a_{it-1}, h(t))$ is the vector of observed states, $\hat{p}$ is a vector of empirical conditional choice probabilities, $\pi(s_{i1})$ is the population probability of type j conditional on initial state s_{i1}, and θ is a vector of the model parameters. The number of unobserved types, J, is assumed to be known and we set $J = 3$. Let $l(a_{it}|s_{it}, j, \hat{p}, \theta)$ denote the likelihood contribution of driver i at time t. We can express the likelihood as follows:

$$l(a_{it}|s_{it}, j, \hat{p}, \theta) = \frac{a_{it} e^{[s_{it}+\rho(\tilde{s}(a=1,s_{it})-\tilde{s}(a=0,s_{it}))]'\theta+\rho(\tilde{e}(a=1,s_{it})-\tilde{e}(a=0,s_{it}))} + (1-a_{it})}{1 + e^{[s_{it}+\rho(\tilde{s}(a=1,s_{it})-\tilde{s}(a=0,s_{it}))]'\theta+\rho(\tilde{e}(a=1,s_{it})-\tilde{e}(a=0,s_{it}))}}$$

where $\tilde{s}(a, s)$ and $\tilde{e}(a, s)$ are known functions of the state s, the conditional choice probabilities $P(a|s)$, and the state transition probabilities $F_s(s|a, s)$. We first estimate the empirical counterparts of the conditional choice probabilities and the transition probabilities of the state variables. Next, we initialize $\hat{\theta}, \hat{\pi}_j(s), \hat{q}_{ij}$ for all $\forall i, j$ by estimating the dynamic model without unobserved types, and set $\hat{\pi}_j(s_1)$ equal to $\hat{\pi}_j(w_1, a_0)$. We then update $\hat{\pi}(s), \hat{q}_{ij}$, $\hat{p}(s, j)$ and $\hat{\theta}$ for all i, j based on the EM algorithm. We refer to Appendix 2.B for further details about the estimation procedure.

Model Fit and Estimation Results

Before we present the estimation results, we examine how well our estimated model fits the data. To examine the fit of the model, we focus on the probability of working conditional on the treatment status and other state variables.

Figure 2.13 plots the probability of working by treatment status across hours of the week, days of the week, and hours of the day. We integrate the model predicted probabilities across all states, except for treatment status and time. The model predicted probabilities fit the data counterparts very well.

In addition to the working pattern by treatment status across time, we also examine the fit of the model by treatment status across demographic groups and unobserved types of drivers. As shown in Appendix 2.E.1, our model predicts well the working patterns across these dimensions.

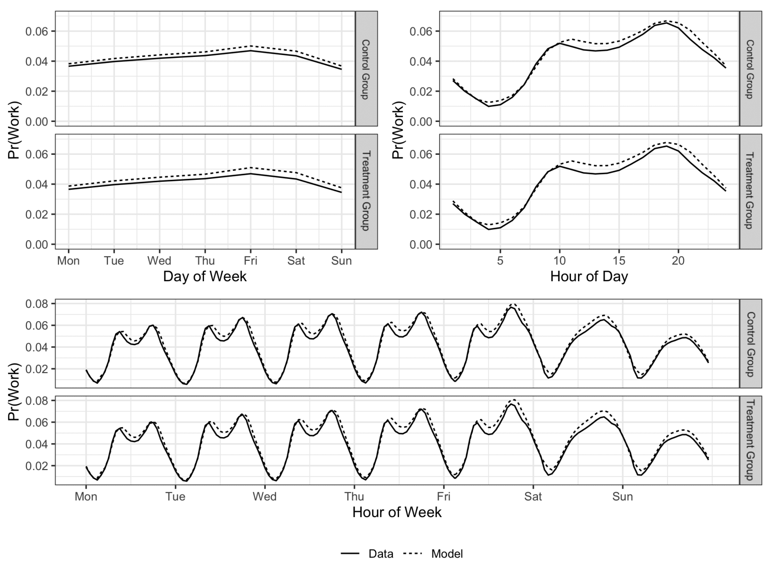


Figure 2.13: Model Fit of Probability of Working by the Treatment Status over Time
Notes: The solid line plots the data, and the dashed line plots the prediction from the model.

Our model also captures the important dynamic component, lagged choices, in a driver's labor supply decision. Figure 2.14 shows the probability of working conditional on the lagged work decision, integrated over the observed heterogeneity and all other states. The probability of working differs significantly depending on whether a driver worked or not in the previous period. Through the lens of the model, this difference produces sizable costs of starting to drive.

The dynamic component is captured by fixed costs of starting to drive in our model. However, one might worry that there are other important sources of adjustment costs. For example, [25] and [26] argue that fatigue is important to understand the behavior of taxi drivers. The reason is that the probability that a taxi driver ends a shift depends strongly on hours worked. In the Uber setting, there is little, if any, evidence of such dependence, as evident from Figure 2.26. Given the weak relationship between the probability of stopping to drive and cumulated hours worked, we decided against including fatigue in the model. Arguably, a small improvement in model fit does not justify to further enlarge the state space, which would significantly increase the computational costs of solving the model.

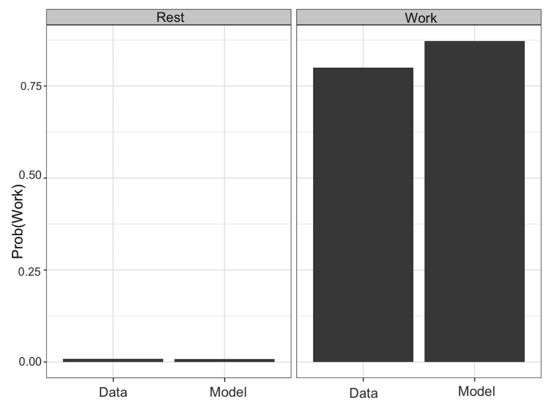


Figure 2.14: Model Fit of Probability of Working by Lagged Work Decisions

Notes: In this figure, the "Rest" panel shows the probability of working conditional on drivers not working in the previous hour. The "Work" panel shows the probability of working conditional on drivers working in the previous hour.

Estimation Results. In Table 2.5, we present the parameter estimates. The estimate of γ, which captures the drivers' sensitivity to wage changes, are positive across the four subgroups. However, the magnitudes vary across the groups. Male drivers are more responsive to exogenous wage changes than female drivers. Consider, for instance, young male drivers. All else being equal, a 1% increase in market wages induces young male drivers to increase their probability of working by around 0.75%.

Our estimation shows considerable dispersion in the value of time $\beta_{h(t)}$ across hours of the week for each of the four subgroups. Recall that a large $\beta_{h(t)}$ in absolute value corresponds to a high cost of driving at $h(t)$, and consequently a low probability of working, all else being equal. On average, the absolute value of the estimated $\beta_{h(t)}$ at the 90th percentile is twice as large as the absolute value of the estimated $\beta_{h(t)}$ at the 10th percentile.

We also find significant adjustment costs in our model as captured by the estimate of μ. For example: Among young male drivers at 8:00 a.m. on Monday, those who worked at 7:00 a.m. have a predicted probability of working around 0.86, while the probability is only 0.07 for those who did not work at 7:00 a.m. This state-dependency emphasizes the importance of incorporating the dynamic component in the driver's decision problem.

Our model allows for three unobserved types among drivers. The type parameter, η, shifts the cutoff in the work choice equation. We normalize η to 0 at all hours of the week for drivers of the baseline type. Thus, high η shifts up the value of work and increases the probability of working relative to the baseline type.

Our estimates of η's suggest three types of drivers. One type is likely to drive in the evening. Another type drives frequently both during the day and at night. The third type, the baseline type, drives infrequently. Evening drivers ($\eta_{1,Night}, \eta_{1,Day}$) have a much higher cutoff of work at night than frequent drivers ($\eta_{2,Night}, \eta_{2,Day}$). Across all the four demographic groups of drivers, the estimated $\eta_{1,Night}$ is almost twice as large as the estimate of $\eta_{2,Night}$. Consider, for example, young male drivers. Conditional on not working in the previous period, the model predicts evening drivers have a probability of working as high as 0.08 at midnight 12:00 a.m., whereas frequent drivers' work probability is only 0.01 and infrequent drivers' work probability is close to zero.

			Weighted Average	Old Male	Young Male	Old Female	Young Female
Preference for Wage	γ		0.036	0.040	0.042	0.024	0.021
Time Preferences	β	$E[\beta_{h(t)}]$	-1.544	-1.612	-1.760	-1.257	-0.977
		$Sd(\beta_{h(t)})$	0.513	0.516	0.462	0.597	0.570
		Median$(\beta_{h(t)})$	-1.328	-1.375	-1.598	-0.972	-0.734
		$q_{10}(\beta_{h(t)})$	-2.295	-2.428	-2.351	-2.164	-1.861
		$q_{90}(\beta_{h(t)})$	-1.055	-1.151	-1.301	-0.653	-0.432
Adjustment Cost	μ		-6.377	-6.191	-6.269	-6.771	-6.864
Unobserved Types	η						
		$\eta_{(1,Night)}$	1.857	1.847	1.901	2.039	1.571
		$\eta_{(1,Day)}$	0.601	0.589	0.669	0.547	0.483
		$\eta_{(2,Night)}$	1.178	1.120	1.237	1.288	1.067
		$\eta_{(2,Day)}$	0.794	0.773	0.802	0.838	0.788
Selection Term	$\frac{\rho_{u\xi}}{\sigma}$		-0.039	-0.043	-0.044	-0.028	-0.024

Table 2.5: Estimates of Model Parameters

Notes: "Weighted average" is calculated by averaging the estimates of the four demographic groups weighted by the share of the drivers. "Young" is defined as those whose ages are less than or equal to the median age.

Table 2.5 also reveals that the estimates of the correction term $\frac{\rho_{u\xi}}{\sigma}$ are negative in all four demographic groups. Recall that $\rho_{u\xi}$ is the correlation coefficient between the preference shock ξ_{it} and the wage component u_{it}. Our estimation results indicate that the costs of working tend to co-move with the market wages. As shown in Appendix 2.C, it is important to take this endogeneity into account to obtain reliable estimates of the preference parameters.

2.5 Insights from the Model

We now use the estimated model to compute the labor supply elasticities, the reservation wages as well as to perform counterfactual analyses. These counterfactuals allow us to infer the drivers' willingness to pay for the ability to customize and adjust their work schedule.[10]

Labor Supply Elasticities

To interpret the magnitude of the preference parameters, we use our estimated model to calculate two types of labor supply elasticities. The first is the Frisch elasticity for the labor supply decision of whether to drive in a given hour of the week. This extensive margin Frisch elasticity can be defined in our model as the percent change in the probability of working in a given hour of the week for an anticipated and temporary one percent exogenous increase in the hourly wage. Formally, we follow [19] and define the Frisch labor supply elasticity on the extensive margin per hour as:

$$
\begin{aligned}
\delta^F \equiv & \frac{\log(\partial Pr(a_{it}=1|s_{it},w_{it}))}{\partial \log(w_{it})} \\
= & \frac{\partial(\Delta V(s_{it},w_{it}) - (\nu_{it} - \epsilon_{0it}))}{\partial w_{it}} f(\Delta V(s_{it},w_{it}) - (\nu_{it} - \epsilon_{0it})) \frac{w_{it}}{Pr(a_{it}=1|s_{it},w_{it})} \\
= & \gamma f(\Delta V(s_{it},w_{it}) - (\nu_{it} - \epsilon_{0it})) \frac{w_{it}}{Pr(a_{it}=1|s_{it},w_{it})}
\end{aligned}
\tag{2.3}
$$

where $\Delta V(s_{it},w_{it}) \equiv V(a_{it}=1,s_{it},w_{it}) - V(a_{it}=0,s_{it},w_{it})$ is the difference between the values of work and rest, $V(\cdot)$ is the value function, s_{it} is the vector of state variables excluding w_{it}, ν_{it} and ϵ_{0it} are idiosyncratic shocks, and f is the probability density function of the choice which follows the logistic distribution. To calculate the elasticity, we first use the

10. Throughout the counterfactual analyses, we abstract from how the market wages may be affected by changes in the labor supply of the drivers. Taking into account such effects would require data on and a model of the demand side of the market.

estimated model parameters to recover the value function V. With the value functions, we then compute the elasticities evaluated for every possible realization of the state variables, using equation (2.3). We average the resulting elasticities weighted by the share of the states in the data and report them in Table 2.6.

Our Frisch labor supply elasticities range from 0.36 to 0.83, with a weighted average of 0.65. On average, male drivers have higher labor supply elasticities than female drivers. Even conditional on observables, there is substantial variation in the elasticity: Infrequent and frequent drivers have the highest elasticities, evening drivers the lowest. By way of comparison, our model implies much smaller differences in the labor supply elasticity over time than across drivers.

		Weighted Average	Daytime	Evening	Weekday	Weekend
		Frisch elasticity				
Observed heterogeneity	Old male	0.71	0.67	0.78	0.68	0.76
	Young male	0.76	0.72	0.83	0.74	0.82
	Old female	0.43	0.41	0.48	0.42	0.47
	Young female	0.38	0.36	0.41	0.36	0.41
Unobserved heterogeneity	Frequent	0.65	0.61	0.74	0.63	0.71
	Evening	0.52	0.54	0.49	0.52	0.54
	Infrequent	0.65	0.62	0.70	0.63	0.70
		IES				
Observed heterogeneity	Old male	0.47	0.42	0.56	0.45	0.50
	Young male	0.52	0.47	0.60	0.50	0.56
	Old female	0.30	0.27	0.37	0.29	0.33
	Young female	0.26	0.23	0.32	0.25	0.28
Unobserved heterogeneity	Frequent	0.35	0.32	0.43	0.34	0.38
	Evening	0.28	0.31	0.22	0.28	0.29
	Infrequent	0.48	0.44	0.58	0.47	0.52

Table 2.6: Model Implied Extensive Margin Labor Supply Elasticities

Notes: "Weighted Average" is calculated by averaging the estimates of the four demographic groups weighted by the share of the drivers. Young is defined as age less than or equal to the median age.

The other elasticity we compute is the intertemporal elasticity of substitution (IES). The IES between two periods, t and $\tilde{t}$, can be defined as follows:

$$\delta_{t,\tilde{t}}^{IES} \equiv \frac{\partial \log(\frac{Pr(a_{it}=1|s_{it},w_{it})}{Pr(a_{i\tilde{t}}=1|s_{it},w_{it})})}{\partial \log(\frac{w_{it}}{w_{i\tilde{t}}})} \qquad \text{for } \tilde{t} > t$$

and measures how much a driver is willing to substitute work between t and $\tilde{t}$ for changes in the relative wage $\frac{w_{it}}{w_{i\tilde{t}}}$. As shown in Appendix 2.E.2, the following expression shows the close link between the IES and the Frisch elasticity:

$$\delta_{t,\tilde{t}}^{IES} = \delta^F - \int \delta_{\tilde{w}}^F g(w_{i\tilde{t}} = \tilde{w}, s_{i\tilde{t}} = \tilde{s}, s_{it}, w_{it}; \mu) d(\tilde{w}, \tilde{s}) \qquad \text{for } \tilde{t} > t \tag{2.4}$$

where $\delta_{\tilde{w}}^F \equiv \frac{\partial \log(P(a_{i\tilde{t}-1}=1|w_{i\tilde{t}}=\tilde{w}, s_{i\tilde{t}}=\tilde{s}, s_{it}, w_{it}))}{\partial \log(w_{it})}$ is the Frisch elasticity of labor supply in $\tilde{t}-1$ if drivers had perfect foresight of all states in $\tilde{t}-1$, and $g(\cdot;\mu)$ is a known increasing function in the adjustment cost μ. The magnitude of the gap between IES and the Frisch elasticity is governed by the adjustment cost μ. When $\mu = 0$, the function $g(\cdot;\mu)$ becomes 0, and hence the Frisch elasticity and the IES coincide.

To calculate IES, we use the estimated model parameters to recover δ^F and $g(\cdot;\mu)$, and then calculate the IES using equation (2.4). The results are presented in Table 2.6. Our IES estimates range from 0.22 to 0.60, with a weighted average of 0.45. The IES estimate is about 30% smaller than the Frisch elasticities, suggesting a considerable size of the adjustment cost.

Value of Time and Reservation Wages

A key objective of the model is to recover the reservation wages and study how they vary over time and across people. We start with presenting the value of time, $\beta_{h(t)}$, which is an important component of the reservation wages.

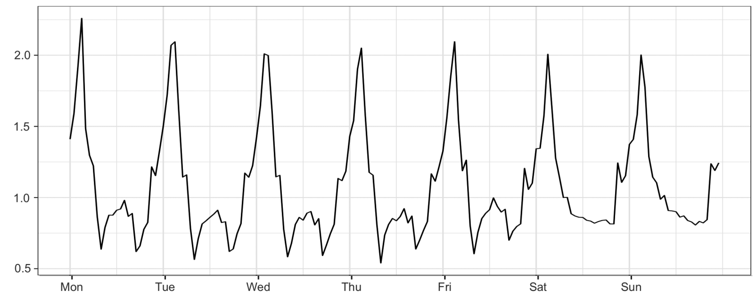


Figure 2.15: Variation in $\beta_{h(t)}$ Relative to Saturday 8 a.m. over Time

Notes: We compute for each hour of the week the weighted average of $\beta_{h(t)}/\beta_{Saturday8a.m.}$ using population shares of each demographic group.

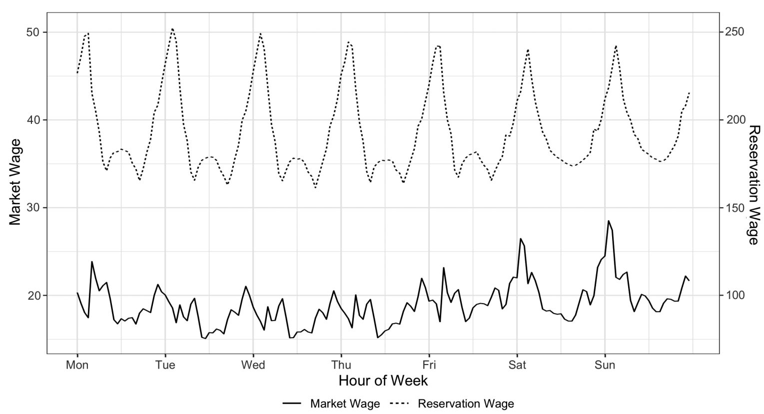


Figure 2.16: Comparison of Observed Market Wages and Reservation Wages

Notes: In this figure, we compute the reservation wage as the minimal wage needed to work given mean preference shocks and averaged across the state variables. The y-axis on the left represents the scale of the market wages, while the y-axis on the right represents the scale of the reservation wages.

Figure 2.15 plots the weighted average of $\beta_{h(t)}/\beta_{\text{Saturday8a.m.}}$ across the four demographic subgroups. Our findings reveal that the value of time varies systematically during a typical week, with the value peaking at late nights around 4 a.m. As shown in Appendix 2.E.3, the value of time varies a lot across hours within a day, whereas there is little variation in the value of time across weekdays. In Figure 2.32 in the Appendix, we find that gender is the key dimension of observable heterogeneity when it comes to the value of time.

Knowledge of the value of time $\beta_{h(t)}$ is necessary but not sufficient to draw inference about reservation wages. We also need to take into account the unobserved preference component ξ_{it}, which may correlate with market wages. In Figure 2.16, we compute and plot the reservation wages against the expected market wages for an average driver during a typical week. The reservation wages stay low during the day, start to increase in the late evening, peak around 4 a.m., and then gradually decline until 9 a.m. The reservation wages are slightly higher during weekends compared to weekdays. While the market wage and the reservation wage tend to co-move across hours, the levels differ significantly. In particular, the average reservation wage is always considerably higher than the average market wage. This finding explains why no more than 4 percent of the drivers are choosing to work in an average hour during the week.

Panel A: Observed Heterogeneity				
	Daytime	Evening	Weekday	Weekend
Old male drivers	1.00	1.28	1.09	1.11
Young male drivers	1.08	1.28	1.14	1.16
Old female drivers	1.93	2.42	2.08	2.11
Young female drivers	2.35	2.88	2.51	2.57
Panel B: Unobserved Types				
	Daytime	Evening	Weekday	Weekend
Frequent drivers	1.00	1.54	1.17	1.20
Evening drivers	1.26	1.12	1.24	1.15
Infrequent drivers	2.33	2.73	2.45	2.49

Table 2.7: Reservation Wages by Driver Types and Time

Notes: In Panel A, the reservation wages are normalized by the mean reservation wage of old male drivers during the daytime. In Panel B, the reservation wages are normalized by the mean reservation wage of frequent drivers during the daytime. The daytime is defined as 6 a.m.-9 p.m., and the evening is defined as 10 p.m.-5 a.m.

The reservation wages vary not only over time but also across people. Panel A of Table 2.7 presents the reservation wages across the four demographic groups and Panel B presents the reservation wages across the three types of drivers. We normalize the reservation wages by the daytime reservation wage of older male drivers. On average, the female drivers' reservation wages are twice as high as the male drivers' reservation wages. Furthermore, young drivers tend to have slightly higher reservation wages than older drivers. There is also a lot of heterogeneity across drivers conditional on age and gender. The frequent drivers have half as large reservation wages as compared to the infrequent drivers, and the evening drivers have lower reservation wages in the evenings as compared to the daytime.

Value of the Ability to Set Customized Work Schedules

Another key objective of the model is to infer the value of job flexibility. One important source of job flexibility is the driver's ability to set a customized work schedule, so she can plan when to work based on her expectations about reservation wages relative to market wages. We quantify the value of this job flexibility by restricting the driver's ability to set the work schedule. In particular, we remove certain hours of a day or days of a week from the choice set of a driver, and then solve the driver's problem given this restricted choice set.

For computational reasons, it is useful to consider a situation where each person drives a given number of consecutive hours per week. In our counterfactual analyses, we let each driver work 5 consecutive hours per week. This choice matches the average number of work hours among the drivers in our estimation sample. At the beginning of a week, each driver is required to choose one 5-hour block to work. We consider two scenarios: The benchmark case and the restricted case. In the benchmark case, drivers can choose among all possible 5-hour blocks. In the restricted case, we remove certain blocks from the choice set. The only difference between the benchmark case and the restricted case is the restriction on the drivers' choice set. Thus, by comparing a driver's utility in these two cases, we can calculate the wage multiplier that she would need to accept a restricted choice set.

In Figure 2.17, we plot the average wage multipliers that the drivers would demand to accept various restrictions on the choice set. In the left panel, we remove the drivers' preferred hours from their choice set. In the calculations behind the first bar, we remove the favorite 5-hour block. The resulting wage multiplier is 1.05. In other words, the average worker would require 5 percent higher wages to accept such a restriction on the choice set. In the second bar, we remove the entire day containing the favorite 5 hours from the driver's choice set. The wage multiplier barely increases. In the third bar, we remove the preferred 5 hours from each day of the week. The wage multiplier now exceeds 1.1.

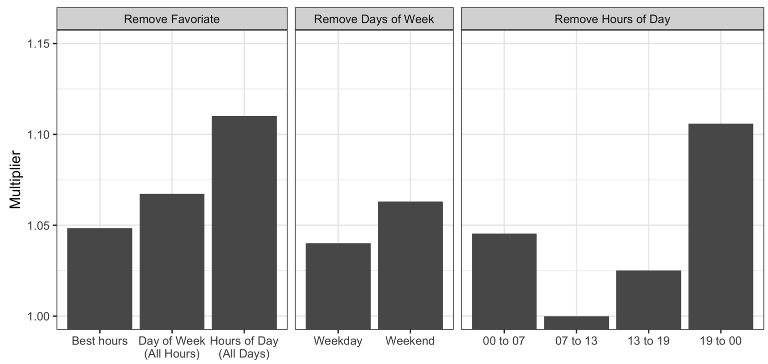


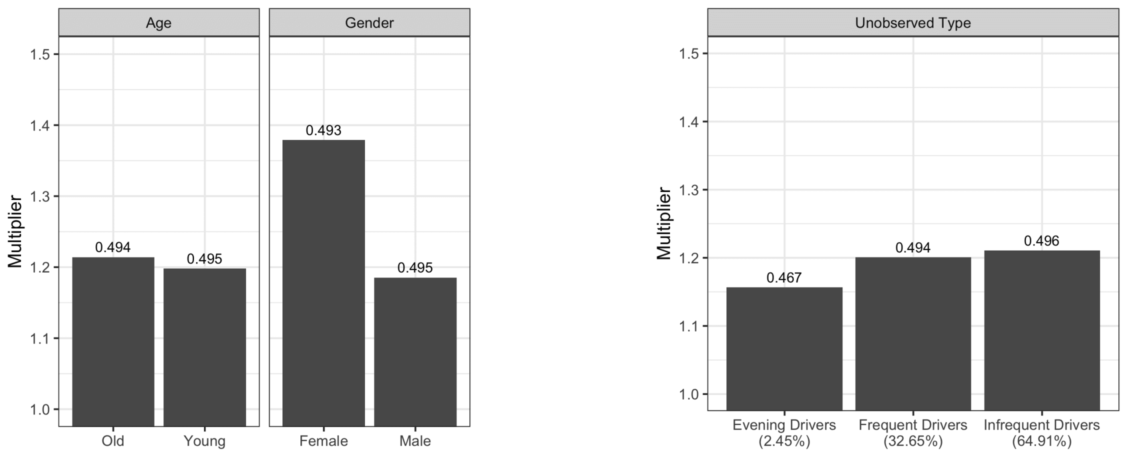
Figure 2.17: Wage Multipliers to Accept Restrictions on the Choice Set

Notes: The left panel removes each drivers' favorite 5-hour block from their choice sets. The "Best hours" indicates the best 5-hour block of the week for the drivers. The middle panel removes weekdays or weekends from drivers' choice sets. The right panel removes certain 5-hour blocks across the entire week from drivers' choice sets.

In the second and the third panel, we restrict the choice set of all drivers to certain days of a week or certain hours of a day. We find that removing the morning block (7 a.m. to 1 p.m.) results in a small wage multiplier. This is because relatively few drivers prefer to work during this time of the day. In contrast, the drivers would demand a large wage multiplier if the evening block (7 p.m. to midnight) would be removed from the choice set. In Figure 2.34 in the Appendix, we show that young drivers and female drivers require relatively high multipliers. This suggests that the ability to set customized work schedules is more important for these groups. We also show in the same figure that frequent drivers and especially evening drivers would require large multipliers to accept restrictions on the possibility of driving in the evening and at night.

Value of the Ability to Adjust Work Schedules

Another source of job flexibility comes from the drivers' ability to adjust work schedules from day to day or even hour to hour in response to unanticipated changes to market wages or the costs of driving. We quantify the value of this job flexibility by forcing the driver to commit to a work schedule before observing the realizations of the innovations to wages and preferences. In particular, we compare the expected values of the two distinct types of work arrangements. The first is a commitment scheme where each driver commits to working the 5-hour block that gives her the highest expected utility. The alternative is a flexible scheme in which each driver is allowed to adjust their choice of 5-hour block once she observes the realizations of the innovations to wages and preferences. By comparing a driver's utility in these two cases, we can calculate the wage multiplier that she would need to accept the commitment scheme instead of the flexible scheme.



(a) Demographic Groups (b) Unobserved Types

Figure 2.18: Wage Multipliers to Accept the Commitment Instead of the Flexible Scheme

Notes: In Figure (a), we compute the weighted average of the wage multipliers needed for drivers in each demographic group to be indifferent between the flexible and the commitment scheme. In Figure (b), we compute the wage multipliers needed for drivers of each unobserved type to be indifferent between the flexible and the commitment scheme. The estimates above the bars show the fraction of drivers that will switch to the second-best 5-hour block.

In Figure 2.18, we report the average wage multipliers that the drivers would demand to accept the commitment scheme. For now, each driver is only allowed to adjust her work schedule once per week in the flexible scheme. On average, about half of the drivers would use this flexibility and change their work schedule due to preference or wage shocks. The wage multiplier needed for drivers to accept the commitment scheme is 1.21. In other words, the average worker would require 21 percent higher wages to prefer the commitment scheme over the flexible scheme. Female drivers place a higher value on the flexibility of adjustment to shocks than male drivers; they need a multiplier of 1.38, while male drivers only require 1.19.

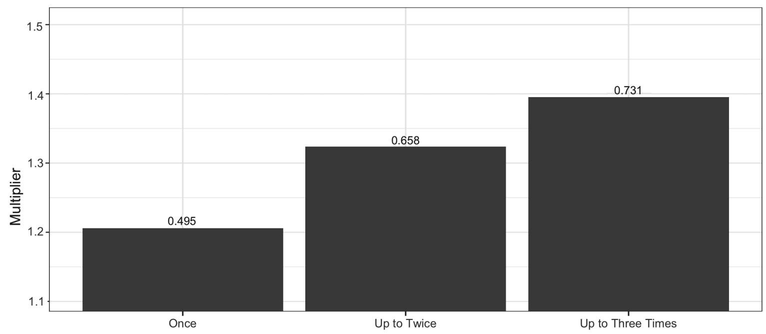


Figure 2.19: Wage Multipliers to Accept the Commitment Scheme Instead of the Flexible Schemes Where Drivers Can Adjust Work Schedules Once, Twice or Three Times

Notes: We compute the weighted average of the wage multipliers needed for drivers to be indifferent between the commitment scheme and the three types of flexible schemes where drivers are allowed to adjust to shocks once, twice, or three times. The estimates above the bars show the fraction of drivers that will switch away from the first best window.

Next, we examine how the wage multipliers change as we increase the number of times drivers are allowed to adjust to shocks. In this analysis, we compare the commitment scheme to flexible schemes where drivers are allowed to adjust once, or twice, or three times within the next week. The results are presented in Figure 2.19. The wage multiplier approaches 1.4 when we compare the commitment scheme to a flexible scheme where each driver can make three adjustments to her weekly work schedule as she observes the realizations of her wage and preference shocks.

Ability to Both Customize and Adjust Work Schedules

In Figure 2.35 and 2.36, we quantify the value of the ability to both customize and adjust work schedules. To this end, we consider three types of work schedules: (1) a fully fixed work schedule where drivers are assigned a particular work schedule and cannot adjust, (2) a commitment scheme where workers can customize but not adjust work schedules, and (3) a flexible scheme where workers can both customize and adjust work schedules. Comparing (3) to (1) allows us to infer the total value of the ability to both customize and adjust work schedules, while the comparison between (3) and (2) isolates the value of the ability to adjust work schedules.

In Figure 2.35, we show that the total value of the ability to both customize and adjust work schedules corresponds to a 43 percent increase in wages. About half of this wage increase can be attributed to the value of the ability to adjust work schedules. Figure 2.36 reveals that female and evening drivers place a particularly high value on work flexibility. In Figure 2.37, we further examine how the wage multipliers change as we increase the number of times drivers are allowed to adjust in (3). The wage multiplier exceeds 1.45 when we compare the commitment scheme to a flexible scheme where each driver can make many adjustments to her weekly work schedule as she observes the realizations of her wage and preference shocks.

Substitution Possibility to Lyft

In the period and cities we consider, Uber has a large majority of the U.S. consumer ride-sharing market. Nevertheless, one might be worried that our estimates are affected by substitution between Lyft and Uber.

To assess this, we perform two sets of analyses. First, we examine if our findings differ systematically in cities or in periods in which Lyft has relatively high or low market share. Second, we compare the results for all drivers to those we obtain from a subsample of drivers who are ineligible to drive for Lyft. The results are reported in Appendix 2.D. In each set of analyses, we find that neither the experimental estimates nor the estimated model parameters depend materially on the possibility of substitution between Lyft and Uber.

2.6 Conclusion

Over the years labor markets have varied dramatically in both their flexibility for where and when the agent works. Just considering the U.S. in the past 75 years, World War II pushed more than 9 million women who previously worked at home to enter factory and other office positions. Work life was a fixed daily and hourly routine completing inflexible tasks. The past decade has witnessed a shift back, with the Gig Economy providing much higher levels of flexibility, autonomy, and task variety. We leverage a natural field experiment at Uber to quantify how reservation wages and labor supply elasticities vary between people and over time, and to infer workers' valuation of flexibility in their choice of work hours. Economists and policymakers are keenly interested in these quantities, especially lately with the growth in jobs that offer flexible work schedules. Combining the experiment with high frequency panel data on wages and work decisions, we documented how labor supply responds to exogenous changes in offered wages in a setting with no restrictions on hours choices stemming from the demand side of the market. We found evidence of systematic heterogeneity in labor supply responses between people and over time, significant fixed costs to starting to drive, and high demand when it is costly for drivers to work.

These experimental findings motivated a dynamic model of labor supply with flexible heterogeneity in preferences over work schedules, start up costs, and the correlation between offered wages and costs of driving in a given period. The primitives of the model were recovered from a combination of the experimental estimates and other data moments. We used the estimated model to compute how labor supply elasticities and reservation wages vary between people and over time, and to perform counterfactual analyses. These analyses allowed us to infer drivers' willingness to pay for the ability to customize and adjust their work schedule.

2.A Wage Prediction Model

We now describe how we construct the expected market wages. As described in Section 2.3, a driver's expected hourly wage is the product of two components: a wage multiplier and the pre-multiplier wage. In the subsections below, we describe each component in detail.

2.A.1 Wage Multipliers

The wage multiplier augments a driver's baseline compensation for driving. It is the maximum of the surge level and the GSL experimental wage multiplier, i.e. max(GSL, Surge). However, since both GSL and surge levels vary over time and across places, this complicates how we construct the wage multipliers.

We illustrate how GSLs vary at the time×place level and how surge levels could limit the experimental variation in GSLs with Table 2.8. The table provides a simple example of a GSL menu for drivers in Chicago which was divided into five regions during most of our sample period: North, West, Loop, South, and Others.

Across all five regions, wherever the GSL experiment is switched on, the treated drivers receive 0.1 higher GSL multipliers than the control group drivers. However, the treatment effect on the wage multiplier is not always 0.1 for two reasons. First, the wage multiplier takes on the maximum value of GSL and the surge level for a given region and time. In regions where the surge levels are higher than the treatment group's GSL level, all drivers in the region receive the same wage multiplier (i.e., the surge level) regardless of their treatment status. Second, there is spatial variation in GSLs and surge levels so GSLs may be higher than surge levels in some regions but not in the others. As a result, we compute the average of the maximum of GSL and surge levels for each driver, weight them by the driver's fraction of time driving in each region in the past, and use this weighted average to predict the driver-specific wage multiplier at every hour.

	Region in Chicago				
	South	North	Loop	West	Others
Treatment	1.5×	1.3×	1.0×	1.1×	1.0×
Control	1.4×	1.2×	1.0×	1.0×	1.0×

Table 2.8: GSL Example in Chicago Across Regions

Notes: Within each region, the treatment group has a GSL multiplier 0.1× higher than the control group.

2.A.2 Pre-multiplier Wages

We calculate the pre-multiplier wages as the observed hourly wages divided by the calculated wage multipliers. We then fit the following regression model to the panel data on the pre-multiplier wages:

$$\tilde{W}_{it} = \alpha_i + \kappa_{h(t)} + \epsilon_{it}$$

where t is an hour, $h(t)$ is the hour of the week at t, i denotes a driver, α_i and $\kappa_{h(t)}$ are driver and hour-of-week fixed effects, and $\tilde{W}$ is the pre-multiplier hourly wage. For each hour t, we fit the model with the panel data up to $t-1$, and use the estimated fixed effects to compute a predicted value for $\tilde{W}$ for every worker at every hour. The predicted hourly wage, $\hat{w}_{it}$, is then constructed as the product of the wage multiplier and the predicted pre-multiplier wage.

2.A.3 Other Prediction Model and Cross-Validation

To assess how well our prediction model performs, we compare it to alternative approaches using a cross-validation procedure. This procedure repeatedly divides samples into a training sample and a testing sample. For each approach, we use the training samples to estimate a prediction model, and then we use the estimated model to form a prediction and calculate the out-of-sample mean squared errors (MSE) on the testing samples. Formally, MSE is defined as the squared distance between the predicted hourly wage rate, $\hat{w}_{it}$, and the observed hourly wage rate w_{it}:

$$MSE = \frac{1}{N_1} \sum_{t=1}^{T} \sum_{i}^{N} (\hat{w}_{it} - w_{it})^2$$

where N_1 is the number of observed hourly wage w_{it}. The smaller the MSE is, the better the prediction model performs.

In addition to our current prediction model, we consider several alternative approaches with matching models. In these models, we consider two sets of observables for matching. The first set of observables consists of demographic variables, including gender and age. The second set of observables includes drivers' past driving histories.

For each set of observables in the matching models, we consider two ways to form the prediction. The first way uses the mean of the observed hourly wages in the previous week. For example, we predict young male drivers' hourly wages at 9 a.m. on a Monday by their mean hourly wages on the Monday 9 a.m. in the previous week. The second way of prediction uses the rolling average of the observed hourly wages. For example, we take average of young male drivers' observed hourly wages at 9 a.m. on all past Mondays to form the wage prediction.

One challenge in this procedure is that a driver's driving history is high-dimensional. Therefore, we conduct dimensionality reduction on driving histories. We consider two types of drivers' work histories, the histories on observed wages and the histories on minutes worked. We first obtain a 168-dimensional vector for each driver by averaging observed wages (or minutes worked across) the past weeks. Each element in the 168-dimensional vector corresponds to a driver's average wages (or minutes worked) in an hour of the week. Next, we cluster drivers based on these 168-dimensional vectors by the K-means clustering algorithm. Formally, let $a_{it} = \frac{1}{t} \sum_{s=1}^{t} a_{is}$ be a vector representing the average observed wages (or minutes worked) for each hour of the week up to t for driver i. Driver groups are

Specification			**MSE**
Fixed Effect Model			
Fixed Effects			0.067
Matching Model			
Var. Matched On		*Prediction From*	
Age × Gender		Last Week	0.077
Age × Gender × History of Wage		Up to Last Week	0.088
Age × Gender × History of Wage		Last Week	0.067
Age × Gender × History of Minutes Worked		Up to Last Week	0.077
Age × Gender × History of Minutes Worked		Last Week	0.088

Table 2.9: Cross-Validation

Notes: MSE refers to the mean squared error of each prediction model. "Prediction From" indicates the periods used to form wage predictions.

obtained by solving the following:

$$arg\min_{k_1,...,k_N} \sum_{i=1}^{N} ||a_{it-1} - \bar{a}_{kt-1}||_2$$

where $i \in \{1, ..., N\}$ denotes a driver, $k_i \in \{1, ..., K\}$ denotes driver i's group, and $\bar{a}_{k_i t-1}$ denotes the mean of a_{it-1} of the group k_i.

We present the results in Table 2.9. The demographic groups we consider include age and female dummies. The driving history includes drivers' histories on observed hourly wages and histories on minutes worked. Our fixed effect wage prediction model outperforms most matching models in terms of out-of-sample mean squared errors. Given its performance and simplicity, we adopt the fixed effect model as our preferred wage prediction model.

2.B Model Identification and Estimation

2.B.1 Model Identification

By Assumption 1, we can rewrite ξ_{it} as

$$\xi_{it} = \frac{\rho_{u\xi}}{\sigma} u_{it} + \psi_{it} = \frac{\rho_{u\xi}}{\sigma}(w_{it} - \delta_{h(t)} - \delta_0 m_t - \delta_1 m_t \times z_{it}) + \psi_{it}$$

where $\psi_{it} \sim \mathcal{N}(0, 1 - \rho_{u\xi}^2)$ and $\psi_{it} \perp u_{it}$ by construction. We define a new state variable $\phi_{it} \equiv w_{it} - \delta_{h(t)} - \delta_0 m_t - \delta_1 m_t \times z_{it}$. Thus, the flow payoffs of the problem become

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \eta_{j(i),h(t)} + \frac{\rho_{u\xi}}{\sigma}\phi_{it} + \nu_{it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

where $\nu_{it} = \psi_{it} + \epsilon_{1it}$. By further substituting w_{it} in the flow payoff with the wage equation, we have

$$U_{it} = \left\{ \begin{array}{ll} \gamma\delta_0 m_t + \gamma\delta_1 m_t \times z_{it} + (\beta_{h(t)} + \gamma\delta_{h(t)}) + \mu \mathbf{1}\{a_{it-1} = 0\} & , \text{ if } a_{it} = 1 \\ \qquad + \eta_{j(i),h(t)} + (\frac{\rho_{u\xi}}{\sigma} + \gamma) u_{it} + \nu_{it} & \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

It is useful to observe that the state variables can be either expressed as $(w_t, h(t), a_{it-1}, \phi_{it}, j(i))$, or equivalently, $(m_t, z_{it}, h(t), a_{it-1}, u_{it}, j(i))$. Note also that we can recover u_{it} from the wage equation, so u_{it} becomes an observed state in the driver's decision problem. Under Assumption 2, the only unobservables in the flow payoffs are ν_{it}, ϵ_{0it}, and $\eta_{j(i),h(t)}$. In the case where unobserved types $j(i)$ were observed, we can identify $(\gamma\delta_0, \beta_{h(t)} + \gamma\delta_{h(t)}, \mu, \frac{\rho_{u\xi}}{\sigma} + \gamma)$ given the observed state $(m_t, z_{it}, h(t), a_{it-1}, u_{it}, j(i))$.

Since $\delta_{h(t)}$, δ_0 and δ_1 are identified from the wage equation, the structural parameters $(\gamma, \beta_{h(t)}, \mu, \frac{\rho_{u\xi}}{\sigma})$ are thus identified. In the case where $j(i)$ is unobserved, we apply (REACH) and (NTYPE) of Assumption 2 as in [37] to jointly identify $\eta_{j(i),h(t)}$ and $(\gamma, \beta_{h(t)}, \mu, \frac{\rho_{u\xi}}{\sigma})$.

2.B.2 Model Estimation

Wage Equation Estimation. We start with estimating the following wage equation using OLS to obtain $\hat{u}_{it}$ and the estimates of $\delta_{h(t)}, \delta_0, \delta_1$:

$$w_{it} = \delta_{h(t)} + \delta_0 m_t + \delta_1 m_t \times z_{it} + u_{it}.$$

Estimation of State Transition. Next, we estimate the joint distribution of the state variables. Recall that we assume the state transition follows a first order Markov process, where the state vector contains hours of the week $h(t)$, treatment status z_{it}, wage component u_{it}, experiment hours m_t, lagged action a_{it-1}, and unobserved type $j(i)$. We also assume that drivers with different unobserved types share the same state transition probability.

Ideally, we can discretize the state space[11] and nonparametrically estimate the transition probabilities. However, the experiment menus vary across blocks and cities, resulting in around 260 distinct experiment menus. Including all of the experiment menus in the estimation of our model greatly enlarges the dimension of the state space. For computational feasibility, we group together similar experiment menus for each city by the K-means clustering method. Specifically, for each experiment menu, we construct a vector of binary indicators for whether an hour has the GSL experiment switched on. We then apply the K-means clustering algorithm on the indicator vectors. We arrive at 5 average experiment menus for each block type and each city. These average experiment menus are taken as inputs in the estimation of the model.

11. Since u_{it} is the only continuous component, we discretize the distribution of $\hat{u}_{it}$ into deciles.

Full Model Estimation. We use the EM algorithm to obtain the maximum likelihood estimates based on equation (2.2), where $s_{it} = (u_{it}, z_{it}, a_{it-1}, h(t), m_t)$ is the vector of the observed states, $\hat{p}$ is a vector of empirical conditional choice probabilities, $\pi(s_{i1})$ is the population probability of type j conditional on the initial state s_{i1}. The number of unobserved types, J, is assumed to be known and equal to 3. We assume that $\hat{\pi}_j(s_1) = \hat{\pi}_j(w_1, a_0)$ so that a driver's unobserved type depends on her initial wage and lagged actions observed in the data.

To obtain the starting values in the EM algorithm, we start with initializing ($\eta_{1,Day}$, $\eta_{1,Night}$, $\eta_{2,Day}$, $\eta_{2,Night}$). The initialization of the remaining structural parameters is set to the estimates of a structural model without any unobserved type. We set the tolerance level to be $1e - 7$.

2.C Comparison with Alternative Model Specifications

We illustrate the potential bias that may arise from more restrictive model specifications. We progressively build up from the simplest possible discrete choice model to highlight how several of our modeling choices – including adjustment costs, permanent observed and unobserved heterogeneity, and the field experiment to address wage endogeneity – are key not only to match the data but also for the estimates of the reservation wages and for the results from the counterfactual analyses. The estimates of all the model specifications discussed are reported in Table 2.14.

2.C.1 Static: Constant Time Preference and Exogenous Wage

Setup. At the beginning of each period, a driver observes the wage w_{it} and the idiosyncratic preference shocks $(\epsilon_{0it}, \epsilon_{1it})$, and she chooses between work ($a_{it} = 1$) and rest ($a_{it} = 0$):

$$U_{it} = \begin{cases} \gamma w_{it} + \beta + \epsilon_{1it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{cases}$$

The driver's labor supply decision depends only on the wage, idiosyncratic shocks, and a constant utility of work β. Once the driver observes the realization of the wage and the preference shocks, she chooses to work if and only if $\gamma w_{it} + \beta \geq \epsilon_{0it} - \epsilon_{1it}$. We maintain the assumptions (IID) and (EXOG) as in Section 2.4. We further assume that ϵ_{0it} and ϵ_{1it} are T1EV and independent: $\epsilon_{1it} \perp \epsilon_{0it}$. In this model, we specify the wage equation as $w_{it} = \delta_{h(t)} + u_{it}$. We also show the results of a nonparametric wage process as a robustness check in the later section.

Figure 2.20: Model Fit of Probability of Working Across Hours of the Week

Notes: The solid line plots the data and the dashed line plots the prediction from the model.

Model Fit. Figure 2.20 compares the model implied probability of working across hours of a week against the data counterparts. While the data reveals a clear picture of systematic variation in the probabilities of working over time, the simple static model does not predict enough variation. Since the preference shocks are IID across time and drivers in our model, the variability in the probability of working comes only from the variability in wages. Therefore, we conclude that the variation in wages is not enough to generate the large variation in the probabilities of working observed in the data.

2.C.2 Static: Varying Time Preference and Exogenous Wage

Since the variation in wages is not sufficient to generate the observed variation in the probability of working over time, we allow the utility of work to vary across hours of the week in the model.

Setup. At the beginning of each period, a driver observes w_{it} and the idiosyncratic shocks $(\epsilon_{0it}, \epsilon_{1it})$, and she chooses a_{it}:

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \epsilon_{1it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

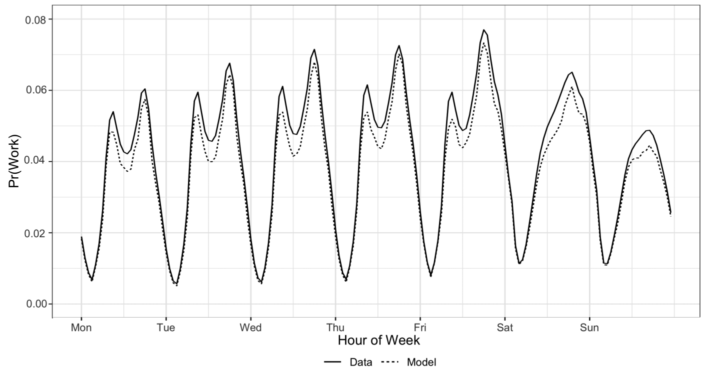
where $h(t)$ is the hour of a week at t, and $\beta_{h(t)}$ is the time-specific shifter of the cost of driving at a given hour of the week. We maintain the above assumptions and the parametric specification on the wage process.

Model Fit. Figure 2.21a compares the predicted probability of working across hours of a week against the data counterparts. With time-varying time preferences, the model now successfully captures drivers' working patterns across hours of a week.

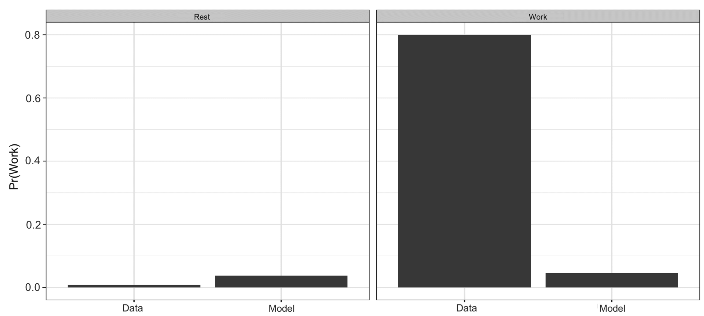
2.C.3 Dynamic:Time-Varying Preference and Exogenous Wage

Although the static model specification in the previous section fits the working patterns across hours of a week excellently, it fails to capture the dynamics of work that we observe in the data. In Figure 2.21b, we compare the probabilities of working conditional on lagged work choices. The static model predicts no difference in the probabilities of working between those who worked in the previous period and those who rested. Yet we observe in the data a much higher probability of working if one worked in the previous period. This state-dependency emphasizes the importance of incorporating the dynamic component in the driver's decision problem to capture the connection between the decision to drive in the current period and future utility. We now consider a dynamic model with adjustments costs and time-varying preferences.

Setup. At the beginning of each period, a forward-looking driver observes w_{it} and the idiosyncratic preference shocks $(\epsilon_{0it}, \epsilon_{1it})$, and forms expectations on future wages and pref-



(a) Hours of the Week



(b) Lagged Choices

Figure 2.21: Model Fit of Probability of Working

Notes: In Figure (a), the solid line plots the data and the dashed line plots the prediction from the model. In Figure (b), the "Rest" panel shows the probability of working conditional on drivers not working in the previous hour. The "Work" panel shows the probability of working conditional on drivers working in the previous hour.

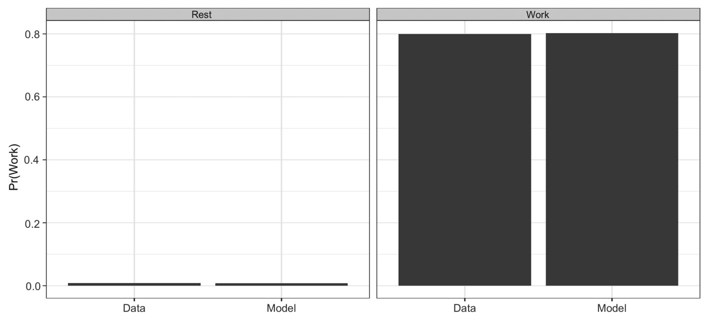
erence shocks. The driver then chooses a_{it} based on the following flow payoffs, taking into account her current choice will affect the expected payoffs in the future:

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \epsilon_{1it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

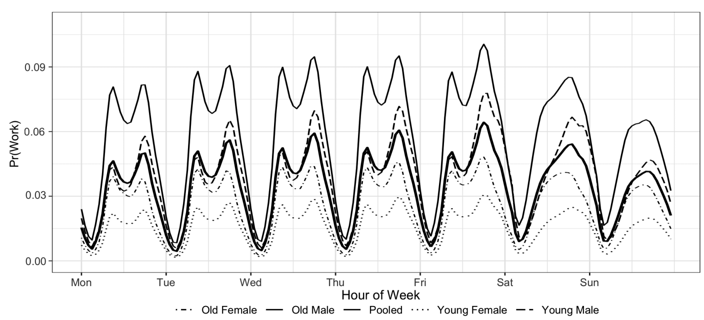
Note that the flow payoff from work depends on whether the driver worked in the previous period; if the driver did not work at $t-1$, she needs to pay an adjustment cost μ to start to work. Both the adjustment cost and the wage transition are linking together periods, since a driver who has paid the adjustment cost in t has a higher continuation value in $t+1$. With fixed costs, the model now becomes a dynamic discrete choice problem. We maintain the above assumptions and the parametric specification of the wage process. In addition, we assume (CI-X) and (DISCOUNT) with the following modification:

(CI-X') State transition probability F satisfies

$$F(w_{it+1}|a_{it}, a_{it-1}, w_{it}, h(t), \epsilon_{1it}) = F(w_{it+1}|a_{it}, a_{it-1}, w_{it}, h(t))$$



(a) Lagged Choices



(b) Demographic Groups

Figure 2.22: Model Fit of Probability of Working

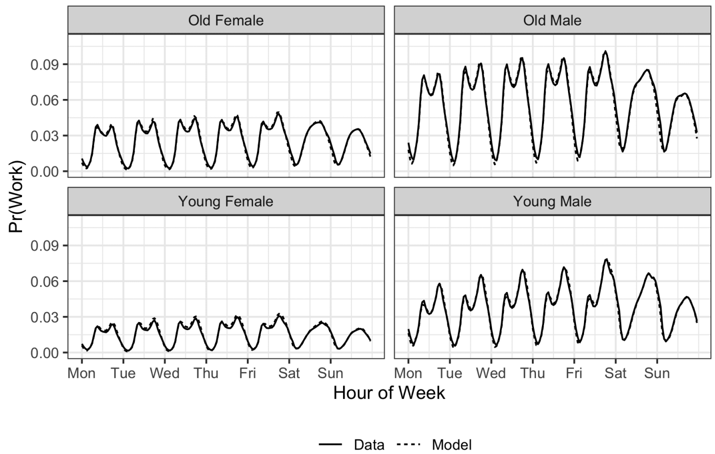
Notes: In Figure (a), the "Rest" panel shows the probability of working conditional on drivers not working in the previous hour. The "Work" panel shows the probability of working conditional on drivers working in the previous hour. In Figure (b), we estimate the probability of working across hours of the week for each demographic group.

Model Fit. Figure 2.22a plots the actual and model implied probabilities of working conditional on the lagged work choice. The dynamic model with the adjustment cost fits the probabilities of working conditional on the lagged choice very well. We, therefore, conclude that (some form of) adjustment costs need to be included to meaningfully represent the driver's decision problem.

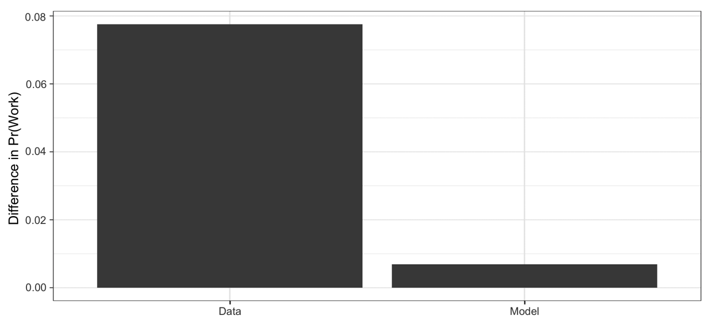
2.C.4 Observed Heterogeneity

So far, we have assumed no systematic preference heterogeneity among drivers. Conditional on wages, the hour of a week, and the lagged choice, the only source that generates differences in drivers' choices is the idiosyncratic preference shocks. Figure 2.22b plots the probability of working across the four subgroups of drivers partitioned by gender and age. This figure shows substantial heterogeneity across the four demographic groups in the probabilities of working. On average, male drivers and older drivers work more than other drivers. This large difference in the probabilities of working across the demographic groups suggests that it is key to let preferences of the drivers to vary by observables such as gender and age.

Estimation and Model Fit. In order to capture the observed heterogeneity along the demographic (gender $\times$ age) dimension, we estimate the dynamic model subgroup by subgroup. The dynamic model fits excellently each demographic group's working patterns. Figure 2.23a plots the conditional choice probabilities for each subgroup against the data counterparts.



(a) Hours of the Week Across Demographic Groups



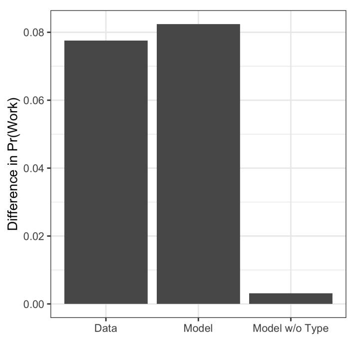
(b) Bottom 50 Percent v.s. Top 50 Percent

Figure 2.23: Model Fit of Probability of Working

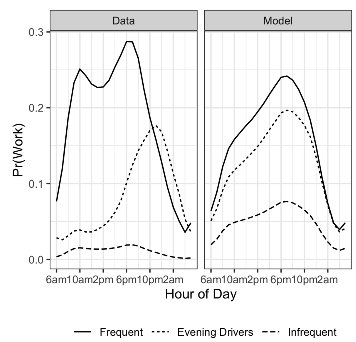
Notes: In Figure (a), the solid line plots the data and the dashed line plots the prediction from the model. We estimate the probability of working across hours of a week for each demographic group in the data and compare it with the probability predicted by the model. In Figure (b), we estimate the difference in the average probability of working between drivers in the top 50 percent and those in the bottom 50 percent, where drivers are ranked by the average hours worked per week.

2.C.5 Time-Invariant Unobserved Heterogeneity

We have shown that there exists a lot of observed heterogeneity across drivers. However, there may also be important unobserved heterogeneity across drivers in the costs or non-wage benefits of working. To illustrate this, Figure 2.23b shows the difference in the average probabilities of working of what we refer to as full-time drivers and infrequent drivers. Full-time drivers have average hours worked above the median. Infrequent drivers have average hours worked below the median. The rank of average hours worked is calculated by pooling all drivers from the four subgroups. As is shown in Figure 2.23b, there is a large difference in the average work probabilities between the top 50 percent and the bottom 50 percent of the drivers, pointing out that there is a great deal of heterogeneity even conditional on observables. We now introduce persistence in the costs or non-wage benefits of working to the model by including unobserved driver types in the flow payoff of work.



(a) Bottom v.s. Top 50 Percent



(b) Hours of the Day by Types

Figure 2.24: Model Fit of Probability of Working

Notes: In Figure (a), we estimate the difference in the average probability of working between drivers above and below the median in the distribution of the average hours worked per week. We conduct the estimation on the data, the simulated data from the model without unobserved types, and the simulated data from the model with two time-invariant unobserved types. In Figure (b), we plot the average probabilities of working within a day of three groups of drivers: (1) infrequent drivers whose average hours worked in the daytime and in the nighttime both rank in the bottom 80 percent, (2) evening drivers whose average hours worked in the daytime rank in the bottom 80 percentile but hours worked in the nighttime rank in the top 20 percent, and (3) the remaining frequent drivers. The ranks are calculated by pooling all four observed subgroups. The daytime is defined to be 6 a.m.-9 p.m., and the nighttime to be 10 p.m.-5 a.m. next day.

Setup of Time-Invariant Unobserved Types. At the beginning of each period, a forward-looking driver observes w_{it} and the idiosyncratic shocks $(\epsilon_{0it}, \epsilon_{1it})$, and forms expectations on future wages and shocks. She then chooses a_{it} based on the following:

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \eta_{j(i)} + \epsilon_{1it} & \text{, if } a_{it} = 1 \\ \epsilon_{0it} & \text{, if } a_{it} = 0 \end{array} \right\}$$

where $j(i)$ is driver i's unobserved type, and $\eta_{j(i)}$ is a shifter of the cost of driving, e.g., full-time drivers have higher $\eta_{j(i)}$ and are thus more likely to drive. We maintain all the assumptions above, use the parametric specification of the wage equation, and invoke the assumption (REACH) and (NTYPE). For now, we assume there are two unobserved types of drivers, full-time drivers and infrequent drivers.

Model Fit. Figure 2.24a plots the difference in the average probabilities of working between drivers whose average hours worked rank above the median and those below the median. There is a significant improvement in the model fit once we include the two unobserved types to the model. Capturing such unobserved persistence in driving is important to model the labor supply of Uber drivers.

2.C.6 Time-Varying Unobserved Types

The assumption of time-invariant unobserved types implies that observationally equivalent workers may systematically differ in how much they work, but not in when they work a lot. In the Uber setting, however, there is likely to be a subset of drivers who work mostly in the evenings due to daytime jobs other than driving for Uber. To illustrate this, Figure 2.24b plots the average probabilities of working within a day of three groups of drivers: (1) infrequent drivers whose average hours worked in the daytime and in the nighttime both rank among the bottom 80 percent,[12] (2) evening drivers whose average hours worked in the daytime rank among the bottom 80 percent but hours worked in the nighttime rank among the top 20 percent, and (3) the remaining frequent drivers. Figure 2.24b reveals a large

12. We define the daytime to be 6 a.m.-9 p.m. and the nighttime to be 10 p.m.-5 a.m. next day.

amount of heterogeneity in the persistence of working within a day across the three types of drivers. By comparison, the model with time-invariant unobserved types is not able to generate such time-specific persistence in driving, even conditional on observables. Motivated by this finding, we introduce time-varying unobserved types to the previous model.

Setup of Time-Varying Unobserved Types. At the beginning of each period, a forward-looking driver observes w_{it} and the idiosyncratic shocks $(\epsilon_{0it}, \epsilon_{1it})$, and forms expectations on future wages and shocks. She then chooses a_{it} based on the following:

$$U_{it} = \left\{ \begin{array}{ll} \gamma w_{it} + \beta_{h(t)} + \mu \mathbf{1}\{a_{it-1} = 0\} + \eta_{j(i),h(t)} + \epsilon_{1it} & , \text{ if } a_{it} = 1 \\ \epsilon_{0it} & , \text{ if } a_{it} = 0 \end{array} \right\}$$

where $\eta_{j(i),h(t)}$ is a shifter in the cost of driving for driver i with unobserved type j at the hour of week $h(t)$. For example, evening drivers have lower $\eta_{j(i),h(t)}$ for $t \in$ Morning but higher $\eta_{j(i),h(t)}$ for $t \in$ Evening, so they are more likely to drive only in the evenings. We maintain all the assumptions above and assume there are 3 unobserved driver types.

Model Fit. Figure 2.25 shows the average probabilities of working of the three groups of drivers (infrequent drivers, full-time drivers, and evening drivers) in the data and in the simulated data from the model. The model captures reasonably well the three unobserved types of drivers after taking into account the observed heterogeneity. The patterns of work vary distinctly over time across the three types of drivers. This is true both in the actual and the simulated data.

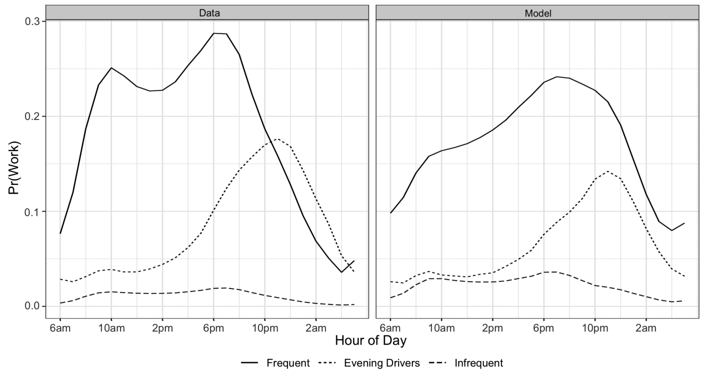


Figure 2.25: Model Fit:Work Probability Across Hours of the Day by Driver Types

Notes: We plot the average probabilities of working within a day of three groups of drivers: (1) infrequent drivers whose average hours worked in the daytime and in the nighttime both rank in the bottom 80 percent, (2) evening drivers whose average hours worked in the daytime rankin the bottom 80 percent but hours worked in the nighttime rank in the top 20 percent, and (3) the remaining frequent drivers. The ranks are calculated by pooling all four observed subgroups. The daytime is defined to be 6 a.m.-9 p.m., and the nighttime to be 10 p.m.-5 a.m. next day.

2.C.7 Probability of Working by Cumulated Hours of Work

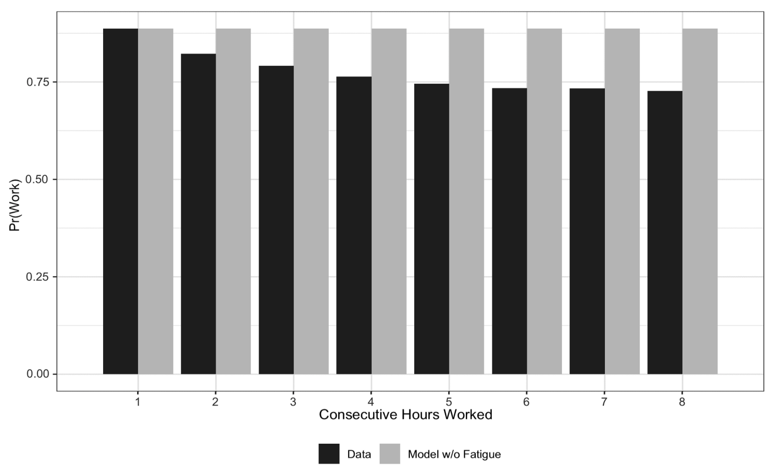


Figure 2.26: Model Fit of Probability of Working by Consecutive Hours Worked

Notes: We define consecutive hours worked as the total number of hours a driver works in a row. The consecutive hour resets to zero whenever a driver stops driving in an hour.

2.C.8 Model of Wages

Before allowing for endogeneity of wages, we show that our estimation results from the previous model do not materially change if we maintain the assumption of exogenous wages but relax the restriction on the process for how wages evolve over time. We relax the wage equation by allowing for a nonparametric evolution of wages as follows. The wage transition is assumed to follow a first-order Markov process. The wage distribution nonparametrically depends on lagged wages, lagged choices, and the hour of the week.[13] We compare the parameter estimates from the models with and without the parametric assumptions on wages in Table 2.10. The estimates of the structural parameters are very similar between the types of models.

Until now, we have assumed exogeneity of wages, ruling out the possibility that market wages may co-move with the cost of driving. As we point out in Section **??**, the downward bias in the OLS estimates of labor supply responses to wage changes is consistent with demand being high when it is costly or difficult to drive. The inclusion of fixed effects of workers and hours of a week helps reduce but not eliminate the bias, suggesting that market wages and the cost of driving may co-move due to idiosyncratic factors, such as weather conditions or entertainment events. This motivates the use of the experiment to address concerns about wage endogeneity.

13. We first discretize the wage distribution into decile grids, and then nonparametrically estimate the transition probability of the wage process conditional on lagged wages, lagged choices, and the hour of the week.

			Non-parametric Wage Model				Parametric Wage Model			
			Old Male	**Young Male**	**Old Female**	**Young Female**	**Old Male**	**Young Male**	**Old Female**	**Young Female**
Preference for Wage	γ		0.004	0.007	-0.009	-0.006	0.004	0.007	-0.009	-0.005
Time Preferences	β	$E[\beta_{h(t)}]$	-0.666	-0.757	-0.643	-0.724	-0.673	-0.761	-0.645	-0.751
		$Sd(\beta_{h(t)})$	0.615	0.506	0.797	0.694	0.614	0.506	0.794	0.697
		Median$(\beta_{h(t)})$	-0.400	-0.619	-0.258	-0.427	-0.410	-0.616	-0.282	-0.453
		$q_{10}(\beta_{h(t)})$	-1.642	-1.417	-1.825	-1.793	-1.641	-1.429	-1.840	-1.822
		$q_{90}(\beta_{h(t)})$	-0.093	-0.205	0.052	-0.066	-0.100	-0.208	0.042	-0.091
Adjustment Cost	μ		-5.515	-5.651	-6.303	-6.256	-5.517	-5.653	-6.306	-6.212
Unobserved Types	η	$\eta_{(1,Night)}$	1.313	1.130	1.715	1.513	1.312	1.131	1.713	1.529
		$\eta_{(1,Day)}$	-0.074	-0.009	-0.047	0.007	-0.074	-0.010	-0.500	0.006
		$\eta_{(2,Night)}$	1.213	1.106	1.683	1.575	1.213	1.106	1.678	1.581
		$\eta_{(2,Day)}$	0.339	0.407	0.457	0.498	0.339	0.406	0.458	0.502

Table 2.10: Estimates of Non-Parametric and Parametric Wage Models

Notes: ”Young” is defined as those whose ages are less than or equal to the median age.

			Time-Variant Unobs. Types					Full Model with Control Function				
			Weighted Average	**Old Male**	**Young Male**	**Old Female**	**Young Female**	**Weighted Average**	**Old Male**	**Young Male**	**Old Female**	**Young Female**
Preference for Wage	γ		0.002	0.004	0.007	-0.009	-0.006	**0.036**	**0.040**	**0.042**	**0.024**	**0.021**
Time Preferences	β	$E[\beta_{h(t)}]$	-0.704	-0.666	-0.757	-0.643	-0.724	-1.544	-1.612	-1.760	-1.257	-0.977
		$Sd(\beta_{h(t)})$	0.607	0.615	0.506	0.797	0.694	0.513	0.516	0.462	0.597	0.570
		Median($\beta_{h(t)}$)	-0.468	-0.400	-0.619	-0.258	-0.427	-1.328	-1.375	-1.598	-0.972	-0.734
		$q_{10}(\beta_{h(t)})$	-1.599	-1.642	-1.417	-1.825	-1.793	-2.295	-2.428	-2.351	-2.164	-1.861
		$q_{90}(\beta_{h(t)})$	-0.113	-0.093	-0.205	0.052	-0.066	-1.055	-1.151	-1.301	-0.653	-0.432
Adjustment Cost	μ		-5.757	-5.515	-5.651	-6.303	-6.256	-6.377	-6.191	-6.269	-6.771	-6.864
Unobserved Types	η											
		$\eta_{(1,Night)}$	1.320	1.313	1.130	1.715	1.513	1.857	1.847	1.901	2.039	1.571
		$\eta_{(1,Day)}$	-0.036	-0.074	-0.009	-0.047	0.007	0.601	0.589	0.669	0.547	0.483
		$\eta_{(2,Night)}$	1.277	1.213	1.106	1.683	1.575	1.178	1.120	1.237	1.288	1.067
		$\eta_{(2,Day)}$	0.399	0.339	0.407	0.457	0.498	0.794	0.773	0.802	0.838	0.788
Selection Term	$\frac{\rho_{u\xi}}{\sigma}$							**-0.039**	**-0.043**	**-0.044**	**-0.028**	**-0.024**

Table 2.11: Estimates of the Exogenous Wage Model and the Full Model

Notes: "Weighted average" is calculated by averaging the estimates of the four demographic groups weighted by the share of the drivers. "Young" is defined as those whose ages are less than or equal to the median age.

Table 2.11 compares the point estimates of the model assuming exogenous wages and the one allowing for endogenous wages. Notably, the point estimates of the preference for wage parameter, γ, which captures drivers' sensitivity to wage changes, increase substantially for all of the four subgroups. In the case of female drivers, the sign of γ estimates flips and becomes positive. The estimate of the correction term $\frac{\rho_{u\xi}}{\sigma}$ is negative in all four subgroups. Recall that $\rho_{u\xi}$ is the correlation coefficient between the preference shock ξ_{it} and the wage component u_{it}. Our estimation results indicate that the costs of working tend to co-move with the market wages. Therefore, it is important to take this endogeneity into account to obtain reliable estimates of the preference parameters.

2.C.9 Economic Implications of the Alternative Modeling Choices

We have demonstrated that the fit of the models improves as we gradually build up from the simple static model up to the full specification. The inclusion of the key components in our model also affects the estimates of the key parameters of interest. We now examine how these parameter estimates, such as preference heterogeneity and adjustment costs, influence the implied reservation wages.

	Reservation Wage			
	Daytime	Evening	Weekday	Weekend
Constant Time Preference	1.80	1.80	1.80	1.80
Varying Time Preference	0.81	0.99	0.87	0.88
Adjustment Cost	6.10	7.99	6.70	6.81
Obs. Heterogeneity	3.16	4.18	3.48	3.56
Time-Invariant Unobs. Types	4.87	6.25	5.31	5.38
Obs. & Unobs. Heterogeneity	3.27	4.64	3.70	3.79
Full Model	1.34	1.65	1.43	1.46

Table 2.12: Comparison of Counterfactual Analyses by Model Specifications

Notes: "Reservation Wage" is calculated as the average minimal wages required to work. We normalize the reservation wages to the old male drivers' reservation wage in the daytime in the full model.

In Column 1-4 in Table 2.12, we report the model implied reservation wages (normalized by old males' reservation wages in the daytime). Clearly, the model specification matters for the estimates of reservation wages. Under the static model with a constant preference for time, the model implied expected reservation wages do not vary over time. Once we allow for time-varying preferences, the expected reservation wages exhibits variation over time. Compared with the full model, the exogenous wage models have much higher reservation wages.

2.D Robustness Checks: Substitution to Lyft

We now present the results from two sets of robustness checks. First, we examine if our findings differ systematically in cities or in periods in which Lyft has relatively high or low market shares. Second, we compare the results for all drivers to those we obtain from a subsample of drivers who are ineligible to drive for Lyft.

2.D.1 Variation in Lyft Share Across Markets

Experimental Estimates by Market. Our sample consists of 3 cities (Boston, Chicago, and San Fransisco), each of which experienced differential growth in the market share of Lyft during the sample period. We estimate wage elasticities for each city-month pair, using the GSL experimental variation in wages, and present the elasticities against Lyft market share in Figure 2.27. As the figure suggests, there is no evidence for correlation between wage elasticities and Lyft market shares.

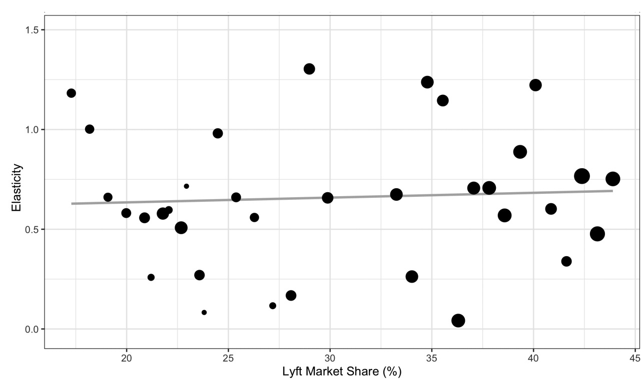


Figure 2.27: Wage Elasticities by Lyft Market Share

Notes: Each dot represents a city × month wage elasticity estimate. The x-axis denotes the Lyft market share. The size of each dot indicates the number of observations in each city × month. The solid line plots the linear regression fit weighted by the number of observations in city × month.

Structural Model Estimates by Market. We now show how much the estimated model parameters vary with the Lyft market share over time and across cities. For each city×time pair,[14] we estimate the model and present the estimates of the value of time parameter, $\beta_{h(t)}$, in Figure 2.28.[15] We see limited changes in $\beta_{h(t)}$ estimates across markets for each demographic subgroup.

2.D.2 Variation in Lyft Eligibility

Experimental Estimates with Lyft Ineligible Drivers. Drivers whose vehicle's model year falls between 2001 and 2003 can only drive for Uber but not for Lyft. In our sample, 3 percent of drivers have vehicles in this range. We define these drivers as Lyft ineligible drivers and estimate the labor supply elasticities on this Lyft ineligible sample. Similar to the labor supply elasticity estimate 0.53 from the sample of all drivers, the Lyft ineligible drivers' elasticity is estimated to be 0.56.

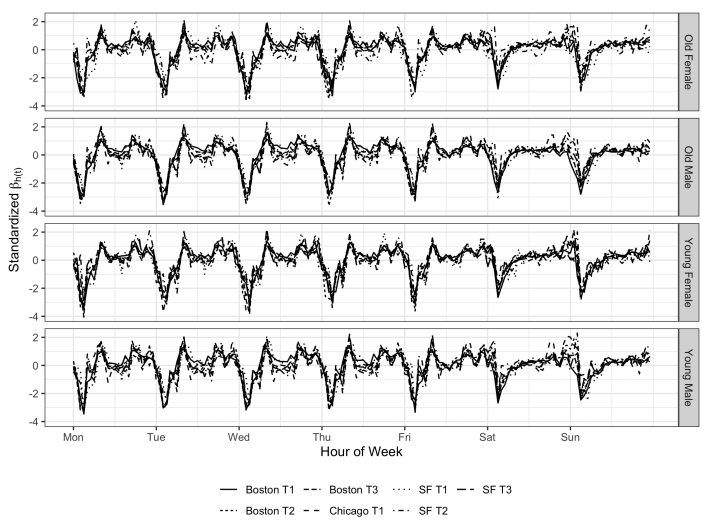


Figure 2.28: Standardized $\beta_{h(t)}$ by Markets and Demographic Groups

Notes: We define a market as a 5-month ×city pair. We estimate the model for each market. T1 represents the period from 2016/10/21 to 2017/03/31. T2 represents the period from 2017/04/01 to 2017/09/31. T3 represents the period from 2017/10/01 to 2018/03/01. We standardize $\beta_{h(t)}$ by de-meaning the raw estimates and dividing them by the standard deviation.

14. We define time as a 5-month period.

15. To avoid empty cells due to sample size restrictions, we estimate a model with endogenous wage but no unobserved heterogeneity.

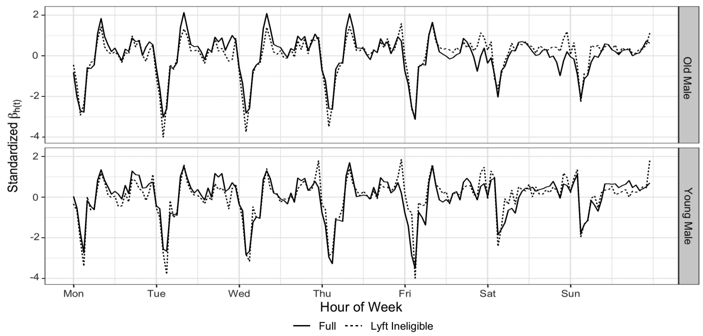


Figure 2.29: Standardized $\beta_{h(t)}$ by Lyft Eligibility

Notes: We standardize $\beta_{h(t)}$ by de-meaning the raw estimates and dividing them by the standard deviation.

Structural Model Estimates with Lyft Ineligible Drivers. Taking advantage of the variation in the Lyft driving eligibility, we separately estimate the structural model on the full sample and the Lyft ineligible sample for male drivers. We allow for heterogeneity by age but no unobserved types, due to the much smaller sample size of Lyft ineligible drivers. In Figure 2.29, we compare the parameter estimates of $\beta_{h(t)}$ for all drivers to those we obtain from the subsample of drivers who are ineligible to drive for Lyft. The estimated value of time parameters, $\beta_{h(t)}$, for Lyft ineligible drivers exhibit a very similar pattern as compared to all drivers. We report the structural parameter estimates from the full sample and the Lyft ineligible sample in Table 2.13. Reassuringly, the structural estimates from the two samples mirror each other.

			Full		Lyft Ineligible	
			Old Male	**Young Male**	**Old Male**	**Young Male**
Preference for Wage	γ		0.042	0.039	0.044	0.043
Time Preferences	β	$E[\beta_{h(t)}]$	-1.152	-0.986	-1.140	-1.082
		$Sd(\beta_{h(t)})$	0.418	0.409	0.592	0.596
		Median$(\beta_{h(t)})$	-1.068	-0.872	-0.995	-0.976
		$q_{10}(\beta_{h(t)})$	-1.641	-1.515	-1.776	-1.691
		$q_{90}(\beta_{h(t)})$	-0.754	-0.607	-0.633	-0.521
Adjustment Cost	μ		-6.479	-7.115	-7.331	-8.176
Selection Term	$\frac{\rho_{u\xi}}{\sigma}$		-0.034	-0.030	-0.037	-0.035

Table 2.13: Estimates by Lyft Eligibility

Notes: "Full" refers to the full sample.

2.E Additional Results on Model

2.E.1 Additional Model Fit Results

Figure 2.30 shows that the working pattern by treatment status across the four demographic groups is well captured by our model, and Figure 2.31 shows that the three unobserved types of drivers are captured reasonably well after taking into account the observed heterogeneity. We use the observed data and the simulated data to plot the average work probabilities within a day of three groups of drivers: (1) infrequent drivers whose average hours worked during the day and at night both rank in the bottom 80 percent,[16] (2) evening drivers whose average hours worked in daytime rank in bottom 80 percent but the hours worked in the

16. We define the daytime to be 6 a.m.-9 p.m., and the nighttime to be 10 p.m.-5 a.m. next day.

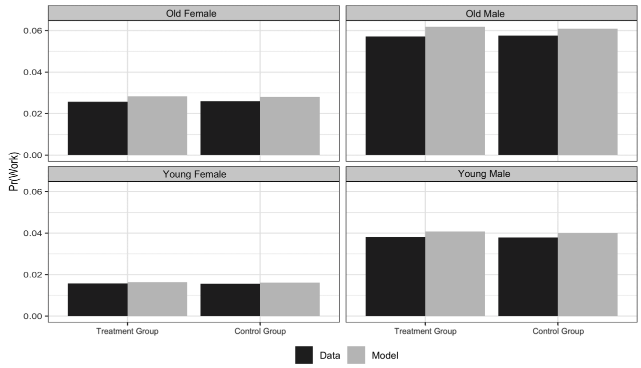


Figure 2.30: Model Fit of Probability of Working by the Treatment Status Across Demographic Groups

Notes: The black bar plots the data, and the grey bar plots the prediction from the model.

nighttime rank in the top 20 percent, and (3) the remaining drivers, who we refer to as frequent drivers because they tend to work a considerable amount both during the day and at night. The patterns of work vary distinctly over time across the three types of drivers. This is true both in the actual and the simulated data.

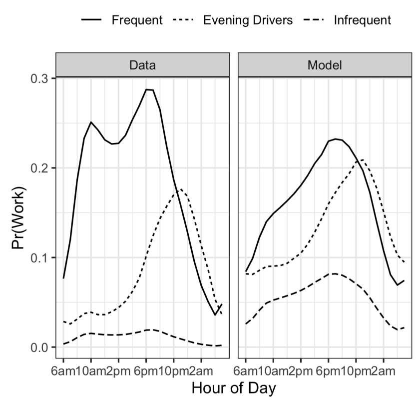


Figure 2.31: Model Fit of Probability of Working by Worker Types Across Hours of the Day

Notes: In this figure, we plot the average work probabilities within a day of three groups of drivers: (1) infrequent drivers whose average hours worked in the daytime and in the nighttime both rank in the bottom 80 percent, (2) evening drivers whose average hours worked in the daytime rank in the bottom 80 percent, but hours worked in the nighttime rank in the top 20 percent, and (3) the remaining frequent drivers. The ranks are calculated by pooling drivers from all four observed subgroups. The daytime is defined to be 6 a.m.-9 p.m., and the nighttime to be 10 p.m.-5 a.m. next day.

2.E.2 Derivation of IES

To derive the desired expression of the IES, we first denote $Pr(a_{it} = 1|s_{it} = s, w_{it} = w)$ by P and $Pr(a_{i\tilde{t}} = 1|s_{it} = s, w_{it} = w)$ by $\tilde{P}$. Denote $\frac{\partial P}{\partial w_{it}}$ by P' and $\frac{\partial \tilde{P}}{\partial w_{it}}$ by $\tilde{P}'$. We set $\tilde{t}=$ t+1 and consider the labor supply response by changing w_{it} around $w_{it} = w$ but holding the distribution of future wage $w_{i\tilde{t}} = \tilde{w}$ fixed. Now, we rewrite δ^{IES} as follows:

$$\delta_{t,\tilde{t}}^{EIS} = \left(\frac{P}{\tilde{P}}\right)' w\frac{\tilde{P}}{P} = \frac{P'}{P}w - \frac{\tilde{P}'}{\tilde{P}}w = \delta^F - \frac{\tilde{P}'}{\tilde{P}}w$$

i.e., the IES is the Frisch elasticity minus the elasticity of future labor supply w.r.t. changes in the current wage. Note that $\tilde{P}$ can be expressed as

$$\begin{aligned}
\tilde{P} &= Pr(a_{i\tilde{t}} = 1|s_{it} = s, w_{it} = w) \\
&= \int\int Pr(a_{i\tilde{t}} = 1|s_{i\tilde{t}} = \tilde{s}, w_{i\tilde{t}} = \tilde{w}, s_{it} = s, w_{it} = w) \\
&\qquad\qquad Pr(s_{i\tilde{t}} = \tilde{s}, w_{i\tilde{t}} = \tilde{w}|s_{it} = s, w_{it} = w)d\tilde{s}d\tilde{w} \\
&= \int\int Pr(a_{i\tilde{t}} = 1|s_{i\tilde{t}} = \tilde{s}, w_{i\tilde{t}} = \tilde{w})Pr(s_{i\tilde{t}} = \tilde{s}, w_{i\tilde{t}} = \tilde{w}|s_{it} = s, w_{it} = w)d\tilde{s}d\tilde{w} \\
&= \int\int Pr(a_{i\tilde{t}} = 1|s_{i\tilde{t}} = \tilde{s}, w_{i\tilde{t}} = \tilde{w})Pr(s_{i\tilde{t}} = \tilde{s}|w_{i\tilde{t}} = \tilde{w}, s_{it} = s, w_{it} = w) \\
&\qquad\qquad Pr(w_{i\tilde{t}} = \tilde{w}|s_{it} = s, w_{it} = w)d\tilde{s}d\tilde{w}
\end{aligned}$$

The third equality is a result of applying (CI-X). We can then simplify $\frac{\tilde{P}'}{\tilde{P}}w$ as

$$\begin{aligned}
\frac{\tilde{P}'}{\tilde{P}}w = \int\int & Pr(w_{i\tilde{t}} = \tilde{w}, a_{i\tilde{t}-1} = 1|s_{it} = s, w_{it} = w) \times \delta_{\tilde{w}}^F \\
& \times \frac{(Pr(a_{i\tilde{t}} = 1|a_{i\tilde{t}-1} = 1, w_{i\tilde{t}} = \tilde{w}) - Pr(a_{i\tilde{t}} = 1|a_{i\tilde{t}-1} = 0, w_{i\tilde{t}} = \tilde{w}))}{Pr(a_{i\tilde{t}} = 1|s_{it} = s, w_{it} = w)} d\tilde{w}d\tilde{s'}
\end{aligned}$$

where $\tilde{s'}$ represents states other than wage in period $\tilde{t}$, and

$$\delta^F_{\tilde{w}} = \frac{w\frac{\partial}{\partial w}Pr(a_{i\tilde{t}-1} = 1|w_{i\tilde{t}} = \tilde{w}, s_{it} = s, w_{it} = w)}{Pr(s_{i\tilde{t}} = 1|w_{i\tilde{t}} = \tilde{w}, s_{it} = s, w_{it} = w)},$$

and we can define

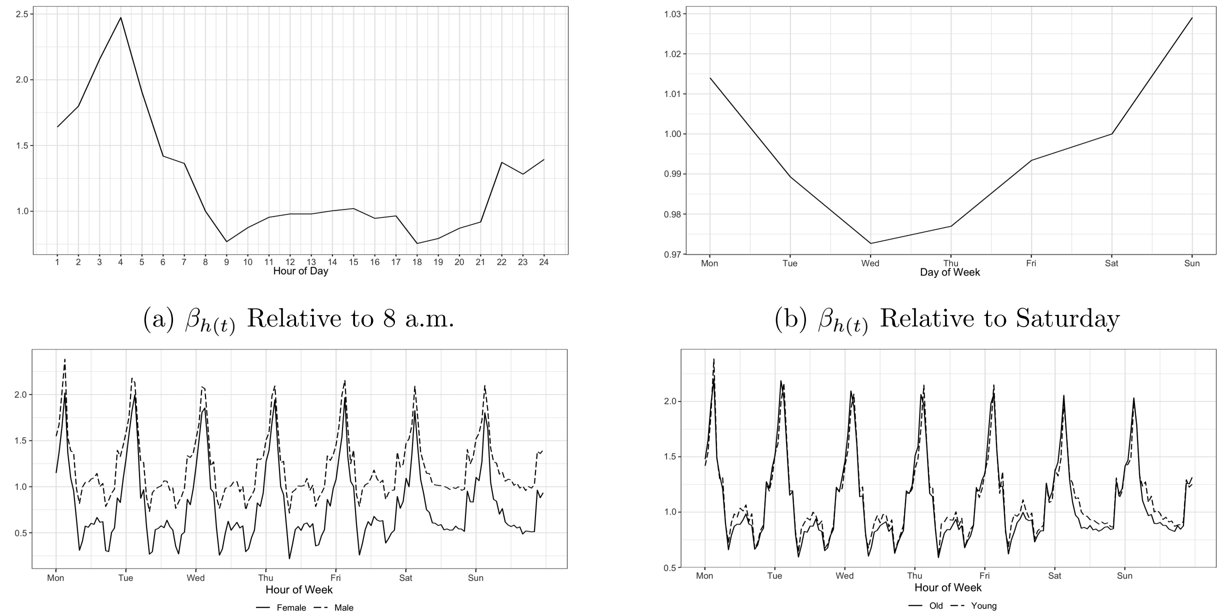
$$g(s, w) \equiv Pr(w_{i\tilde{t}} = \tilde{w}, a_{i\tilde{t}-1} = 1|s_{it} = s, w_{it} = w)$$

$$\times \frac{(Pr(a_{i\tilde{t}} = 1|a_{i\tilde{t}-1} = 1, w_{i\tilde{t}} = \tilde{w}) - Pr(a_{i\tilde{t}} = 1|a_{i\tilde{t}-1} = 0, w_{i\tilde{t}} = \tilde{w}))}{Pr(a_{i\tilde{t}} = 1|s_{it} = s, w_{it} = w)}.$$

2.E.3 Counterfactual: Value of Time and Reservation Wages

To examine whether the variation mostly comes from hours of a day or days of a week, Figure 2.32a shows the average $\beta_{h(t)}$ per hour of the day (relative to 8 a.m.), while Figure 2.32b reports the average $\beta_{h(t)}$ per day of the week (relative to Saturday). The value of time varies a lot across hours within a day. By way of comparison, there is little variation in the value of time across weekdays. The value of time tends to be higher during weekends than weekdays, but the differences are rather small.

In Figure 2.32c and 2.32d , we compare the value of time across the demographic groups. For each group, we normalize the estimates of $\beta_{h(t)}$ by the value of time for older males on Saturdays at 8 a.m.



(a) $\beta_{h(t)}$ Relative to 8 a.m.

(b) $\beta_{h(t)}$ Relative to Saturday

(c) Across Gender

(d) By Age

Figure 2.32: Variation in the Value of Time, $\beta_{h(t)}$,

Notes: In (a), we compute the average $\beta_{h(t)}$ per hour of the day (relative to 8 a.m.) weighted by the shares of each demographic group. In (b), we compute the average $\beta_{h(t)}$ per day of the week (relative to Saturday) weighted by the population shares of each demographic group. In (c) and (d), we compute the weighted average of $\beta_{h(t)}$ relative to old males' value of time on Saturday 8 a.m. using population shares as weights.

2.E.4 Counterfactual: Value of the Ability to Set Customized Work Schedules

Figure 2.33 and 2.34 plot the wage multipliers by gender, age, and driver types.

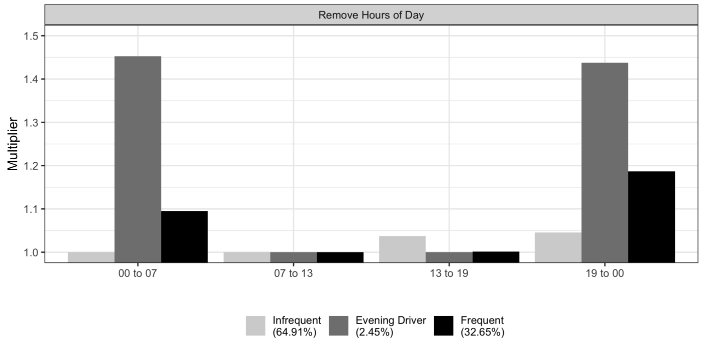
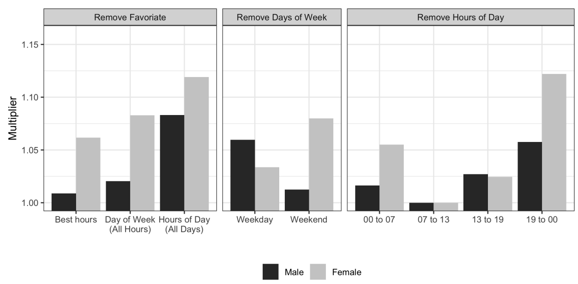
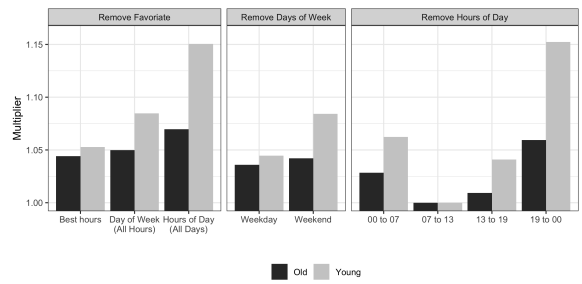


Figure 2.33: Unobserved Types

Notes: We remove certain hour blocks across the entire week from drivers' choice sets. The numbers in the parentheses indicate the fraction of each type of drivers.



(a) Gender



(b) Age

Figure 2.34: Wage Multipliers to Accept Restrictions on the Choice Set

Notes: In Figure (a) and (b), the left panel removes each drivers' favorite 5-hour block from their choice sets. The "Best hours" indicates the favorite 5-hour block of the week for the drivers. The middle panel removes weekdays or weekends from drivers' choice sets. The right panel removes certain hour blocks across the entire week from drivers' choice sets.

2.E.5 Counterfactual: Value of the Ability to Both Customize and Adjust Work Schedules

To infer the value of the ability to both customize and adjust work schedules, we let every driver work 5 consecutive hours per week. For now, we only let the drivers make one adjustment per week in scenario (2) and (3). In scenario (1), we fix the work schedule of each driver to be Friday evening, 7:00 -11:59 p.m.

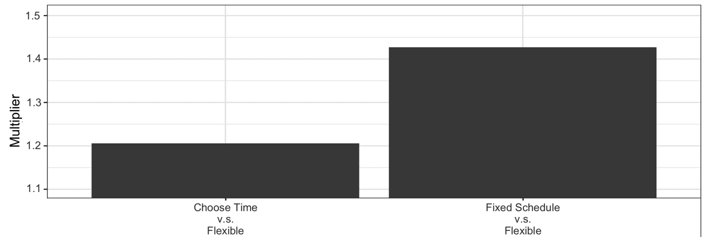
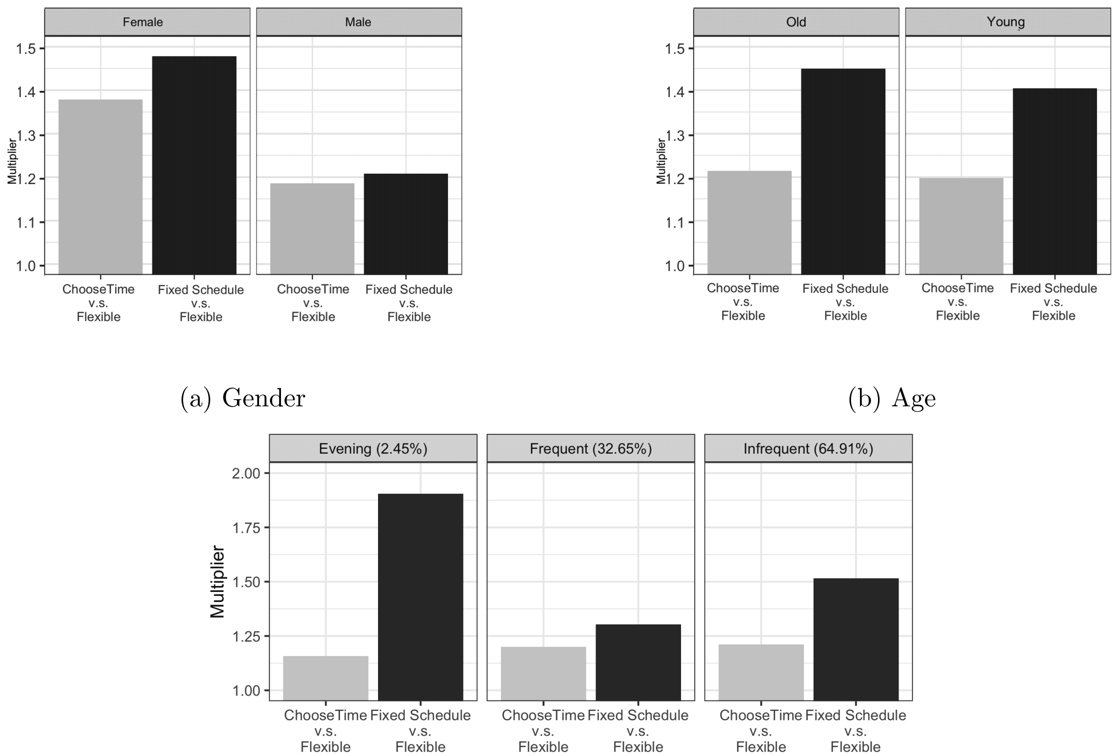


Figure 2.35: Wage Multipliers to Accept Restrictions on the Choice Set and Work Schedule Adjustments

Notes: In this figure, we compute the weighted average of the wage multipliers needed for drivers to be indifferent between the different types of work schedules. The fixed work schedule starts from 7:00 p.m. on Friday and ends at 11:59 p.m. on Friday.



(a) Gender

(b) Age

(c) Unobserved Types

Figure 2.36: Heterogeneity in Wage Multipliers to Accept Restrictions on the Choice Set and Work Schedule Adjustments

Notes: In these figures, we compute the wage multipliers needed for drivers to be indifferent between the different types of work schedules. The fixed work schedule starts from 7:00 p.m. on Friday and ends at 11:59 p.m. on Friday.

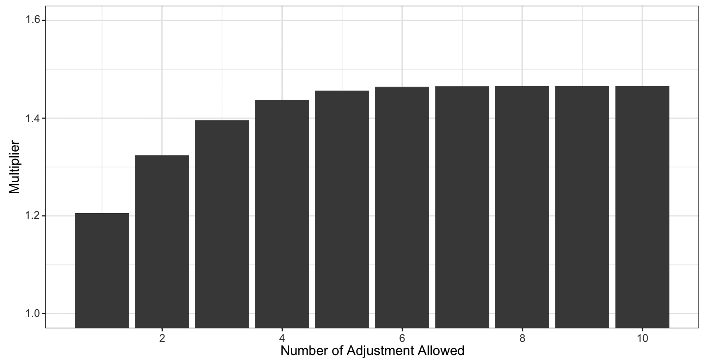


Figure 2.37: Wage Multiplier with Increasing Number of Work Schedule Adjustments

Notes: In this figure, we compute the wage multipliers needed for drivers to be indifferent between the commitment scheme and the flexible scheme with increasing numbers of adjustments.

Specification		Preference for Wage	Time Preferences					Adjustment Cost	Unobserved Types					Selection Term
		γ	β					μ	η					$\frac{\rho_{u\xi}}{\sigma}$
			$E[\beta_{h(t)}]$	$Sd(\beta_{h(t)})$	Median$(\beta_{h(t)})$	$q_{10}(\beta_{h(t)})$	$q_{90}(\beta_{h(t)})$		η	$\eta_{(1,Night)}$	$\eta_{(1,Day)}$	$\eta_{(2,Night)}$	$\eta_{(2,Day)}$	
Constant Time Preference		0.015	-3.692											
Varying Time Preference		0.034	-4.059	0.672	-3.778	-5.259	-3.441							
Adjustment Cost		0.005	-0.376	0.401	-0.283	-0.812	0.021	-6.397						
Obs. Heterogeneity	Weighted Average	0.002	-0.322	0.414	-0.222	-0.790	0.084	-6.453						
	Old Male	0.004	-0.358	0.401	-0.251	-0.881	0.025	-6.018						
	Young Male	0.007	-0.394	0.386	-0.313	-0.751	0.009	-6.470						
	Old Female	-0.008	-0.187	0.489	-0.060	-0.755	0.250	-6.735						
	Young Female	-0.007	-0.129	0.467	-0.024	-0.670	0.320	-7.426						
Time-Invariant Unobs. Types	Weighted Average	0.001	-0.640	0.440	-0.523	-1.161	-0.220	-6.517	0.615					
	Old Male	0.003	-0.660	0.425	-0.540	-1.219	-0.259	-6.148	0.567					
	Young Male	0.007	-0.741	0.409	-0.642	-1.161	-0.320	-6.490	0.612					
	Old Female	-0.013	-0.475	0.527	-0.325	-1.104	-0.057	-6.900	0.719					
	Young Female	-0.009	-0.439	0.493	-0.314	-1.041	0.034	-7.325	0.662					
Time-Variant Unobs. Types	Weighted Average	0.002	-0.712	0.606	-0.477	-1.609	-0.122	-5.753		1.321	-0.037	1.277	0.399	
	Old Male	0.004	-0.673	0.614	-0.410	-1.641	-0.100	-5.517		1.312	-0.074	1.213	0.339	
	Young Male	0.007	-0.761	0.506	-0.616	-1.429	-0.208	-5.653		1.131	-0.010	1.106	0.406	
	Old Female	-0.009	-0.645	0.794	-0.282	-1.840	0.042	-6.306		1.713	-0.050	1.678	0.458	
	Young Female	-0.005	-0.751	0.697	-0.453	-1.822	-0.091	-6.212		1.529	0.006	1.581	0.502	
Time-Variant Unobs. Types (Nonparam)	Weighted Average	0.002	-0.704	0.607	-0.468	-1.599	-0.113	-5.757		1.320	-0.036	1.277	0.399	
	Old Male	0.004	-0.666	0.615	-0.400	-1.642	-0.093	-5.515		1.313	-0.074	1.213	0.339	
	Young Male	0.007	-0.757	0.506	-0.619	-1.417	-0.205	-5.651		1.130	-0.009	1.106	0.407	
	Old Female	-0.009	-0.643	0.797	-0.258	-1.825	0.052	-6.303		1.715	-0.047	1.683	0.457	
	Young Female	-0.006	-0.724	0.694	-0.427	-1.793	-0.066	-6.256		1.513	0.007	1.575	0.498	
Full Model with Control Function	Weighted Average	0.036	-1.544	0.513	-1.328	-2.295	-1.055	-6.377		1.857	0.601	1.178	0.794	-0.039
	Old Male	0.040	-1.612	0.516	-1.375	-2.428	-1.151	-6.191		1.847	0.589	1.120	0.773	-0.043
	Young Male	0.042	-1.760	0.462	-1.598	-2.351	-1.301	-6.269		1.901	0.669	1.237	0.802	-0.044
	Old Female	0.024	-1.257	0.597	-0.972	-2.164	-0.653	-6.771		2.039	0.547	1.288	0.838	-0.028
	Young Female	0.021	-0.977	0.570	-0.734	-1.861	-0.432	-6.864		1.571	0.483	1.067	0.788	-0.024

Table 2.14: Estimates of All Specifications

REFERENCES

[1] John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

[2] Katharine G Abraham and Henry S Farber. Job duration, seniority, and earnings, 1986.

[3] Joseph G Altonji and Robert A Shakotko. Do wages rise with job seniority? *The Review of Economic Studies*, 54(3):437–459, 1987.

[4] Joseph G Altonji and Nicolas Williams. Do wages rise with job seniority? a reassessment. *Ilr Review*, 58(3):370–397, 2005.

[5] Joshua D Angrist, Sydnee Caldwell, and Jonathan V Hall. Uber vs. taxi: A driver's eye view. *American Economic Journal: Applied Economics*, forthcoming.

[6] Neele Balke and Thibaut Lamadon. Productivity shocks, long-term contracts and earnings dynamics. Technical report, National Bureau of Economic Research, 2020.

[7] Erling Barth. Firm-specific seniority and wages. *Journal of Labor Economics*, 15(3):495–506, 1997.

[8] Magali Beffy, Richard Blundell, Antoine Bozio, Guy Laroque, and Maxime To. Labour supply and taxation with restricted choices. *Journal of Econometrics*, 211(1):16–46, 2019.

[9] Richard Blundell and Thomas MaCurdy. Labor supply: A review of alternative approaches. In *Handbook of labor economics*, volume 3, pages 1559–1695. Elsevier, 1999.

[10] Stéphane Bonhomme, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. How much should we trust estimates of firm effects and worker sorting? Technical report, National Bureau of Economic Research, 2020.

[11] Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739, 2019.

[12] Ken Burdett and Melvyn Coles. Equilibrium wage-tenure contracts. *Econometrica*, 71(5):1377–1404, 2003.

[13] David Card, Jörg Heining, and Patrick Kline. Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015, 2013.

[14] M. Keith Chen, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen. The value of flexible work: Evidence from uber drivers. *Journal of Political Economy*, 2019.

[15] Victor Chernozhukov, Iván Fernández-Val, and Ye Luo. The sorted effects method: discovering heterogeneous effects beyond their averages. *Econometrica*, 86(6):1911–1938, 2018.

[16] Raj Chetty, Adam Guren, Day Manoli, and Andrea Weber. Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–75, 2011.

[17] Peter Cohen, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. Using big data to estimate consumer surplus: The case of uber. Technical report, National Bureau of Economic Research, 2016.

[18] Cody Cook, Rebecca Diamond, Jonathan V Hall, John A List, and Paul Oyer. The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. 2020.

[19] John K Dagsvik. Marginal compensated effects and the slutsky equation for discrete choice models. 2020.

[20] Stefano DellaVigna, John A. List, and Ulrike Malmendier. Testing for Altruism and Social Pressure in Charitable Giving *. *The Quarterly Journal of Economics*, 127(1):1–56, 01 2012.

[21] Stefano DellaVigna, John A List, Ulrike Malmendier, and Gautam Rao. Estimating social preferences and gift exchange at work. Working Paper 22043, National Bureau of Economic Research, February 2016.

[22] Christian Dustmann and Costas Meghir. Wages, experience and seniority. *The Review of Economic Studies*, 72(1):77–108, 2005.

[23] Niklas Engbom, Christian Moser, and Jan Sauermann. Firm pay dynamics. Technical report, National Bureau of Economic Research, 2022.

[24] Tor Eriksson and Nicolai Kristensen. Wages or fringes? some evidence on trade-offs and sorting. *Journal of Labor Economics*, 32(4):899–928, 2014.

[25] Henry S Farber. Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of political Economy*, 113(1):46–82, 2005.

[26] Henry S Farber. Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. *The Quarterly Journal of Economics*, 130(4):1975–2026, 2015.

[27] Benjamin Friedrich, Lisa Laun, Costas Meghir, and Luigi Pistaferri. Earnings dynamics and firm-level shocks. Technical report, National Bureau of Economic Research, 2019.

[28] Luigi Guiso, Luigi Pistaferri, and Fabiano Schivardi. Insurance within the firm. *Journal of Political Economy*, 113(5):1054–1087, 2005.

[29] Jonathan V Hall and Alan B Krueger. An analysis of the labor market for uber's driver-partners in the united states. *ILR Review*, 71(3):705–732, 2018.

[30] John Haltiwanger, Henry Hyatt, and Erika McEntarfer. Who moves up the job ladder? *Journal of Labor Economics*, 36(S1):S301–S336, 2018.

[31] Daniel S Hamermesh. Changing inequality in markets for workplace amenities. *The Quarterly Journal of Economics*, 114(4):1085–1123, 1999.

[32] Glenn W Harrison and John A List. Field experiments. *Journal of Economic literature*, 42(4):1009–1055, 2004.

[33] Haoran He, David Neumark, and Qian Weng. Do workers value flexible jobs? a field experiment on compensating differentials. Technical report, National Bureau of Economic Research, 2019.

[34] V Joseph Hotz and Robert A Miller. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529, 1993.

[35] Boyan Jovanovic. Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990, 1979.

[36] Boyan Jovanovic. Matching, turnover, and unemployment. *Journal of political Economy*, 92(1):108–122, 1984.

[37] Hiroyuki Kasahara and Katsumi Shimotsu. Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica*, 77(1):135–175, 2009.

[38] Michael P Keane. Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075, 2011.

[39] Mark R Killingsworth and James J Heckman. Female labor supply: A survey. *Handbook of labor economics*, 1:103–204, 1986.

[40] Patrick Kline, Neviana Petkova, Heidi Williams, and Owen Zidar. Who profits from patents? rent-sharing at innovative firms. *The quarterly journal of economics*, 134(3):1343–1404, 2019.

[41] Marta Lachowska, Alexandre Mas, Raffaele D Saggio, and Stephen A Woodbury. Do firm effects drift? evidence from washington administrative data. Technical report, National Bureau of Economic Research, 2020.

[42] David Laibson and John A. List. Principles of (behavioral) economics. *American Economic Review*, 105(5):385–90, May 2015.

[43] Thibaut Lamadon, Magne Mogstad, and Bradley Setzler. Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review*, 112(1):169–212, 2022.

[44] Nicole Maestas, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger. The value of working conditions in the united states and implications for the structure of wages. Technical report, National Bureau of Economic Research, 2018.

[45] Thierry Magnac and David Thesmar. Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816, 2002.

[46] Isabel Z Martinez, Emmanuel Saez, and Michael Siegenthaler. Intertemporal labor supply substitution? evidence from the swiss income tax holidays. *American Economic Review*, 111(2):506–46, 2021.

[47] Alexandre Mas and Amanda Pallais. Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59, 2017.

[48] Jacob Mincer, Boyan Jovanovic, et al. Labor mobility and wages. *Studies in labor markets*, pages 21–64, 1981.

[49] Derek Neal. Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, 13(4):653–677, 1995.

[50] Daniel Parent. Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics. *Journal of Labor Economics*, 18(2):306–323, 2000.

[51] John Pencavel. Labor supply of men: a survey. *Handbook of labor economics*, 1:3–102, 1986.

[52] Brooks Pierce. Compensation inequality. *The Quarterly Journal of Economics*, 116(4):1493–1525, 2001.

[53] Sherwin Rosen. The theory of equalizing differences. *Handbook of labor economics*, 1:641–692, 1986.

[54] John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

[55] Andy Snell, Pedro Martins, Heiko Stüber, and Jonathan P Thomas. Bias in returns to tenure when firm wages and employment comove: A quantitative assessment and solution. *Journal of Labor Economics*, 36(1):47–74, 2018.

[56] Robert Topel. Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of political Economy*, 99(1):145–176, 1991.

[57] John Van Reenen. The creation and capture of rents: wages and innovation in a panel of uk companies. *The quarterly journal of economics*, 111(1):195–226, 1996.

[58] Matthew Wiswall and Basit Zafar. Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507, 2017.