Hours and Wages in Equilibrium*

Thimo De Schouwer†

January 10, 2025

Abstract

Hours and wages differ substantially across occupations, which has important consequences for inequality. This paper introduces a tractable two-sided search and matching model to explain these patterns and study their consequences. We model equilibrium hours and wages as a function of the preferences of workers, the production technologies of firms, and bargaining power in the labor market. We estimate the model with several sources of data from the United States, and find substantial occupational heterogeneity in the productivity of hours in line with Goldin (2014). We also document an important role for bargaining power, which pushes up wages in occupations where demand exceeds supply.

Preliminary and Incomplete – Please do not Circulate

Keywords: hours and wages, occupations, search and matching, random utility model

JEL Codes: J31, J32, C78

*I thank Edoardo Ciscato and Iris Kesternich for invaluable guidance and support. I am also grateful to Laurens Cherchye, Thibault Deneus, Jeremy Fox, Janoś Gabler, Thomas Le Barbanchon, Tim Mensinger, and Robert Miller, for helpful discussions and feedback, and to audiences at the Household Economics Gathering (Brussels), Applied Microeconomics Workshop (Leuven), BGSE Applied Microeconomics Workshop (Bonn), the Summer School in Labor Economics (IZA), European Association for Labor Economics Conference (EALE, Bergen), and the European Winter Meeting of the Econometric Society (EWMES, Palma). I gratefully acknowledge financial support from the Research Foundation Flanders (FWO) through PhD Fellowship (project grant 11I7822N). All errors are mine.

†Department of Economics, University of Leuven. E-mail: thimo.deschouwer@kuleuven.be.
Website: thimodeschouwer.github.io.

1 Introduction

Hours and wages differ substantially across occupations, which has important implications for inequality (Autor and Dorn, 2013; Goldin, 2014). To explain this pattern, recent work points to non-linearities and heterogeneity in occupational production technologies (Goldin, 2014; Goldin and Katz, 2016; Erosa et al., 2024). But hours and wages are equilibrium outcomes, and also depend on workers' preferences, and on tightness in the labor market. We introduce a tractable equilibrium search and matching model with rich multidimensional heterogeneity to study the importance of all three channels. We estimate the model to look at how equilibrium hours and wages are determined across occupations in the United States, and consider consequences for gender equality.

The model builds on the separable Transferable Utility (TU) framework of Choo and Siow (2006) and Galichon and Salanié (2022). Their model allows for the match surplus to be a non-linear function of the preferences of workers and firms, who differ in observed and unobserved ways. We extend the framework by introducing search frictions, and allow agents to direct their search efforts as in Moen (1997). On the one hand, this allows the model to better rationalize the coexistence of unemployed workers and vacant jobs, which are the consequence of coordination failures instead of voluntary choices (see the discussion in Chade et al. (2017)). On the other, because firms commit to posted wages, the matching problem has an Imperfectly Transferable Utility (ITU) structure, in which market tightness affects equilibrium wages *and* hours (Legros and Newman, 2007; Galichon et al., 2019).

We introduce a static model of the labor market in which equilibrium wages and hours are the outcome of a two stage process. Firms first post job vacancies that specify wages and hours to attract their preferred type of worker. Workers simultaneously apply to their favored positions after observing all postings. The wages that firms offer at different hours reflect heterogeneity in the productivity of hours, but also account for workers' heterogeneous preferences over leisure, and competitively price in the probability of filling the vacancy as in Moen (1997). The number of matches between workers and firms across different numbers of hours depends on the endogenous application decisions of workers, which are in turn a function of the match surplus and tightness in the market.

To estimate the model, we combine several datasets from the United States. The main source of data is the latest American Community Survey (ACS, wave 2022), provided by the Integrated Public Use Microdata Series (IPUMS), from which we take individual level data on wages and hours and compute matching patterns across worker types and occupations. We focus on a sub-sample of college graduates in high earning occupations as in Goldin (2014). The second data source is the National Labor Exchange (NLx) Research Hub's job vacancy database. This is a dataset with over 300 million job postings by firms, and contains information about the occupation and state for each individual vacancy. We use this data to construct occupational job vacancy distributions by state, which we scale using aggregate counts from our third source of data, the Job Openings and Labor Turnover Survey (JOLTS).

The first main result about occupational wages and hours is in line with the narrative of Goldin (2014) and Goldin and Katz (2016) on the importance of occupational heterogeneity production technologies. We find substantial differences in the perceived value of productivity of hours across occupations. Long hours (defined as 50 or more per week) are substantially more productive than full time hours (35 to 50 per week) in Business occupations, but not in Health, Tech, Science, or Others. This translates into large wage premiums and a substantial fraction of employees working long hours in Business occupations but not elsewhere. For part-time work (less than 35 hours per week) we find the opposite patterns, a result that is also in line with the non linear production functions introduced in Goldin (2014).

We also document an important role for bargaining power as a determinant of hours and wages. Particularly in Business occupations, where we find that the number of searching firms substantially exceeds the number of searching workers. The perceived bargaining weights of firms in Business occupations are between 0.30 and 0.40, compared to the average of about 0.44. This pushes up the average wages in these occupations by about 14% compared to the equal weight benchmark. We find the opposite in Science occupations, where firms' bargaining weights are about 0.60, pushing down wages by roughly 6%. We also find heterogeneity in workers' preferences for hours across occupations, but these are less important in shaping equilibrium wages, except in Health occupations.

We then look at how these different channels shape wage inequality between men and women through several counterfactual exercises. We first study what happens when we reduce firms' incentives to reward long hours – as suggested in Goldin (2014). To this end, we remove the productivity premium for working long hours in Business occupations, and study the counterfactual distributions of wages and hours. We also consider a counterfactual where bargaining weights are equalized at 0.50 across occupations.

Outline. The remainder of this paper is organized as follows. Section 2 discusses related literature. Section 3 presents the model. Section 4 introduces the data and presents descriptive statistics. Section 5 discusses estimation. The results are discussed in section 6. Section 7 concludes.

2 Related Literature

Hours and Earnings. This paper is related to the literature that studies the non-linear relation between hours and earnings that started with the early work of Cogan (1981) and French (2005). Recent papers by Goldin (2014) and Goldin and Katz (2016) argue that these non-linearities differ across occupations due to heterogeneity in production technologies leading to compensating differentials as in Rosen (1974, 1986). Erosa et al. (2022) and Erosa et al. (2024) incorporate these ideas into a classic Roy (1951) model to study how occupational differences in hours affect wage and earnings inequality, and Jang and Yum (2022) provides a dynamic life-cycle extension. These papers focus on the sorting of workers faced with exogeneously determined nonlinearities in the wage profile. We instead consider an equilibrium model where the wage distribution is completely endogenous. This means that contrary to these papers we do not need to make prior assumptions about the relation between productivity, hours, and wages. We also allow for richer heterogeneity on both sides of the market and study counterfactual simulations that allow for equilibrium effects.

We also contribute to the substantial body of empirical work that studies non-linearities through penalties for part-time work (Macpherson and Hirsch, 1995; Aaronson and French, 2004; Manning and Petrongolo, 2008) and premiums or penalties for long hours (Kuhn and Lozano, 2008; Goldin, 2014; Yurdagul, 2017; Cortés and Pan, 2019; Mantovani, 2023). Recent work by Bick et al. (2022) summarizes the cross sectional patterns as a hump-shaped relation between hours and wages, with substantial occupational heterogeneity in where wages are maximized. We add to this literature by studying the foundations of these penalties and premiums.

Models of Matching. The model introduced in this paper builds on the large literature that estimates models of (search and) matching with transfers. We build on the approach introduced in Choo and Siow (2006) that introduces additive random preference shocks on both sides of the market. This approach is widely used to explain matching patterns in the marriage market – for example in Dupuy and Galichon (2014), Siow (2015), Chiappori et al. (2017), or Chiappori et al. (2020) – but has seen limited applications to labor markets. There are exceptions in Dupuy et al. (2020), Dupuy and Galichon (2022), Corblet (2023), and Dupuy et al. (2023), but these papers primarily focus on specific high skill and low friction markets. This is mainly because their models abstract from search frictions, which implies that workers are essentially unemployed voluntarily, which is at odds with the large literature that argues for the importance of frictions in the labor market (e.g. the work following Shimer and Smith (2000)). We add to this literature by introducing search frictions into the model of Choo and Siow (2006) and differ in terms of application.

Related work by Jaffe and Weber (2019) and Ciscato (2024) introduces search frictions into Choo and Siow (2006) by allowing for heterogeneous meeting costs between agents of different types.[1] The main difference is that they impose market-clearing conditions at the aggregate level, whereas in our model firms post wages to different disaggregated sub-markets, and workers apply to their single preferred position. Because firms commit to their posted wages, utility is imperfectly transferable (ITU) as in Legros and Newman (2007) and Galichon et al. (2019). This allows for tightness in the sub-markets to affect the endogenous matching terms – hours of work in our model – a channel that is shut down in perfectly transferable utility (TU) models. An alternative way to allow for the market structure to affect the matching terms is discussed in Mourifié and Siow (2021), but comes at the cost of having no clear structural interpretation.

Two other related papers by Arcidiacono et al. (2016) and Beauchamp et al. (2024) introduce directed search in a similar matching model with additive random preference shocks. But in order to separately identify male from female preferences without observing transfers, these papers make the assumption that utility is not transferable (NTU). The implication of doing so is that any concessions between potential match partners are ruled out, which is an unrealistic

[1] See also San (2024) who models social connections in the labor market in a similar fashion.

assumption in our modern labor market setting. Other models of directed search with transfers in the labor market such as Eeckhout and Kircher (2010), Kaas and Kircher (2015), and Schaal (2017) differ in scope, and typically study longitudinal dynamics and conditions for assortative matching, whereas we focus on explaining hours and wages through two-sided multidimensional heterogeneity. Related work by Michelacci and Pijoan-Mas (2012) also studies hours and earnings in a model with search frictions, but focuses on the relation between wage dispersion and long hours. Mantovani (2023) estimates the importance of changes in the returns to long hours over time in a frictionless ITU matching model.

This paper is also related to the literature on hedonic pricing that followed from Rosen (1974) and Ekeland et al. (2004). While we focus on working hours, the model can include and price any other job characteristic. These characteristics can be modeled either as endogenous outcomes or as fixed exogenous firm or occupation characteristics. We extend the classic hedonic pricing framework by simultaneously considering a model of matching in the labor market. This is also a feature of the model by Dupuy (2021), but he abstracts from frictions and focuses on migration decisions in the marriage market.

3 The Search and Matching Model

This section introduces an equilibrium matching model with directed search and endogenous wages and hours.[2] The labor market is populated by workers $i \in I$ and firms $j \in J$ who can match one-to-one.[3] Firms post vacancies that specify the wage, the number of hours, and the type of worker they want to attract. Types are defined based on observable characteristics, and are discrete and multidimensional. We denote workers' types by $x = (1, 2, ..., X)$, firms' types by $y = (1, 2, ..., Y)$, and hour types by $h = (1, 2, ..., H)$. Workers apply simultaneously to their preferred vacancies. Because of coordination failures and imbalances in supply and demand, some jobs remain vacant, and some workers unemployed.

Labor Supply. Workers apply to the job that maximizes their utility. We model the utility of a

[2] While this paper focuses on hours, the model works with any other endogenous job characteristic.

[3] In one-to-one matching models, job and firm are used interchangeably. This is equivalent to a one-to-many matching setting, where several workers may be matched to the same firm *if* the value of production is separable between the different jobs a firm offers (see Roth and Sotomayor (1989)).

worker i of type x searching for a type y job with h hours as the sum of a systematic component and an idiosyncratic preference shock:

$$u_{x_i yh} = m_{\theta_{xyh}}(\alpha_{xyh} + W_{xyh}) + \xi_{x_i yh}. \tag{3.1}$$

The first term of the systematic utility $m_{\theta_{xyh}}$ is the worker's probability of finding a job. We follow Menzio and Shi (2010) and Schaal (2017) and model this job finding probability as a Constant Elasticity of Substitution (CES) function of market tightness θ_{xyh}, which is defined as the ratio of posted vacancies to searching workers:

$$m_{\theta_{xyh}} = (1 + \theta_{xyh}^{\lambda})^{\frac{1}{\lambda}}. \tag{3.2}$$

To assure that this function is concave increasing, such that more firms relative to workers increases the worker's job finding probability, and maps to the unit interval, we assume $\lambda < 0$. The job finding probabilities of workers are multiplied with their valuation of this job, which depends on wages W_{xyh} and the value of amenities offered α_{xyh}. The idiosyncratic shock $\xi_{x_i yh}$ captures individual worker i's preference for jobs of type (y, h), and rationalizes why workers of the same type search for different jobs.

Labor Demand. Firms post the vacancy that maximizes their profits. We also model profits with a systematic component and a shock specific to firm j:

$$\pi_{xy_j h} = q_{\theta_{xyh}}(\rho_{xyh} - W_{xyh}) + \eta_{xy_j h}. \tag{3.3}$$

The first term of the systematic component $q_{\theta_{xyh}}$ represents the firm-side probability of hiring a worker, which we assume to be a decreasing function of market tightness.[4] The probability of hiring a worker is multiplied with the systematic profits that workers generate, which are modeled as their productivity value ρ_{xyh} net of the wage bill W_{xyh}. Note that the productivity of hours is allowed to vary by worker and firm type. We again introduce a shock $\eta_{xy_j h}$ that captures the idiosyncratic preference of firm j over workers of type (x, h). This shock rationalizes why

[4] We follow the literature and define $q_{\theta_{xyh}} = \frac{m_{\theta_{xyh}}}{\theta_{xyh}}$. Concavity of $m_{\theta_{xyh}}$ suffices to ensure that tighter markets reduce the firm's probability of hiring a worker.

firms of the same type search for, and match with, different types of workers.

3.1 Model Solution.

The equilibrium matching and wage functions are determined in two stages. Firms first decide on the type of vacancy to post. This choice incorporates a trade-off between profits and the probability of filling the vacancy. Workers observe all the offers that firms have posted, and then decide where to apply. Their application decision is based on an analogous trade-off between match quality and the probability of finding a job. This is reflected in the equilibrium wages, which are determined as a function of workers' preferences, firms' technologies, and bargaining power as captured through tightness in the labor market.

Equilibrium Wages. We first solve for optimal wages and market tightness. The equilibrium problem can be stated as firms maximizing profits (3.3) subject to workers' utilities (3.1):

$$\begin{aligned} &\max_{\theta,w,x,h} \left\{ q_{\theta_{xyh}}(\rho_{xyh} - W_{xyh}) + \eta_{xy_jh} \right\} \\ &\text{s.t. } u_{x_iyh} = m_{\theta_{xyh}}(\alpha_{xyh} + W_{xyh}) + \xi_{x_iyh}. \end{aligned} \tag{3.4}$$

We substitute out wages and solve for optimal market tightness to recover the equilibrium wage function as in Wright et al. (2021):

$$W_{xyh} = \left(1 - \epsilon_{\theta_{xyh}}\right) \rho_{xyh} - \epsilon_{\theta_{xyh}} \alpha_{xyh}, \tag{3.5}$$

with $\epsilon_{\theta_{xyh}}$ the elasticity of matching with respect to market tightness, defined by:

$$\epsilon_{\theta_{xyh}} := \theta_{xyh} \frac{m'_{\theta_{xyh}}}{m_{\theta_{xyh}}}. \tag{3.6}$$

The equilibrium wage function (3.5) has an intuitive interpretation. Wages increase in productivity ρ_{xyh} and decrease in amenities α_{xyh}. But the extent to which firms reward workers for their increased productivity – or workers sacrifice wages for better amenities – is entirely determined by the elasticity term, which is a function of market tightness. Wages are thus determined as if they were bargained over, with bargaining weights equal to $\epsilon_{\theta_{xyh}} \in (0, 1)$, a decreasing function of market tightness. When there is slack in the labor market, firms face little competition, and $\epsilon_{\theta_{xyh}} \to 1$. Wages are then set to just reward workers for their disutility, as $W_{xyh} \to -\alpha_{xyh}$. When markets become more tight, $\epsilon_{\theta_{xyh}} \to 0$ and workers are paid the entire value of production as $W_{xyh} \to \rho_{xyh}$.[5]

Wage Premia. We can study premiums and penalties to part-time work or long hours within a worker-firm match as:

$$\frac{W_{xyh}}{W_{xyh'}} = \mathcal{P}(\Delta\alpha, \Delta\gamma, \Delta\epsilon), \tag{3.7}$$

where h is a typical full-time position and h' the part-time or longer hour alternative. The premium or penalty associated with deviating from a full time position is a function of the difference in preferences, production technologies, and bargaining weights between the two options. This differs from the frictionless transferable utility benchmark – which we obtain by setting the matching probabilities and market tightness at one. The wage premium is then pinned down only by preferences and productivity.

Equilibrium Matching Patterns. We can substitute the equilibrium wages back into the primitives of the model to obtain the following indirect utility and profit functions:

$$\bar{U}_{x_iyh} = m_{\theta_{xyh}}(1 - \epsilon_{\theta_{xyh}})\Phi_{xyh} + \xi_{x_iyh}, \tag{3.8}$$

$$\bar{\Pi}_{xy_jh} = q_{\theta_{xyh}}\epsilon_{\theta_{xyh}}\Phi_{xyh} + \eta_{xy_jh}, \tag{3.9}$$

where Φ_{xyh} denotes the systematic joint surplus of a match, defined by:

$$\Phi_{xyh} := \alpha_{xyh} + \rho_{xyh}. \tag{3.10}$$

The utility and profits thus increase in the joint surplus Φ_{xyh} and in respectively the fraction captured by workers $(1 - \epsilon(\theta_{xyh})$ and firms $\epsilon(\theta_{xyh})$.

We assume that the preference shocks on both sides of the market are iid Type 1 Extreme Value distributed following Choo and Siow (2006).[6] Then, utility and profit maximization imply

[5] Note that this requires $\rho_{xyh} > -\alpha_{xyh}$, or equivalently $\Phi_{xyh} > 0$, such there are gains to be made from matching. The alternative cannot be sustained as an equilibrium outcome.

[6] This assumption can be relaxed using the tools introduced in Galichon and Salanié (2022).

the following conditional labor supply (S) and demand (D) functions:

$$S_{y,h|x} := \exp\left(m_{\theta_{xyh}}(1-\epsilon_{\theta_{xyh}})\Phi_{xyh}\right) a_x, \tag{3.11}$$
$$\text{where: } a_x = \left(\sum_y \sum_h \exp\left(m_{\theta_{xyh}}(1-\epsilon_{\theta_{xyh}})\Phi_{xyh}\right)\right)^{-1},$$

$$D_{x,h|y} := \exp\left(q_{\theta_{xyh}}\epsilon_{\theta_{xyh}}\Phi_{xyh}\right) b_y, \tag{3.12}$$
$$\text{where: } b_y = \left(\sum_x \sum_h \exp\left(q_{\theta_{xyh}}\epsilon_{\theta_{xyh}}\Phi_{xyh}\right)\right)^{-1}.$$

These functions allow us to recover the equilibrium matching patterns. Let M_{xyh} denote the number of (x, y, h) matches, M_{x0} the number of unemployed workers of type x, and M_{0y} the number of vacant positions of type y. The matching patterns are then defined by:

$$M_{xyh} = m_{\theta_{xyh}} S_{xyh} = q_{\theta_{xyh}} D_{xyh}, \tag{3.13}$$
$$M_{x0} = \left(1 - m_{\theta_{xyh}}\right) S_{xyh}, \tag{3.14}$$
$$M_{0y} = \left(1 - q_{\theta_{xyh}}\right) D_{xyh}. \tag{3.15}$$

Equilibrium Hours. We can use equation (3.13) to study how the equilibrium number of hours is determined in each occupation. Let h again denote a full-time arrangement and h' the part-time alternative. The fraction of workers in full-time as opposed to part-time jobs is then determined by:

$$\frac{M_{xyh}}{M_{xyh'}} = \frac{m_{\theta_{xyh}}}{m_{\theta_{xyh'}}} \exp(m_{\theta_{xyh}}(1-\epsilon_{\theta_{xyh}})\Phi_{xyh} - m_{\theta_{xyh'}}(1-\epsilon_{\theta_{xyh'}})\Phi_{xyh'}) = \mathcal{H}(\Delta\Phi, \Delta\theta), \tag{3.16}$$

with relatively more or less workers in full time positions when $\mathcal{H}(\Delta\Phi, \Delta\theta) \lesseqgtr 1$. Note that the hours function $\mathcal{H}(\Delta\Phi, \Delta\theta)$ takes both the match surplus Φ_{xyh} and market tightness θ_{xyh} as arguments. This is different from the TU model[7], which we can recover by setting $\theta_{xyh} = \theta_{xyh'}$, where the fraction of workers in part-time jobs is then pinned down solely by the difference in the match surplus $\Delta\Phi$, with more workers in full-time positions when $\Phi_{xyh} > \Phi_{xyh'}$.

[7]This property of TU models is extensively discussed in Mourifié and Siow (2021).

Characterizing the Equilibrium. The following adding-up restrictions hold in equilibrium:

$$\begin{aligned} \sum_y \sum_h D_{xyh} &= n_x \text{ for all } x \\ \sum_x \sum_h S_{xyh} &= f_y \text{ for all } y, \end{aligned} \tag{3.17}$$

where n_x denotes the total number of workers of type x and f_y the total number of firms of type y. Together with an equation for market tightness, which can be derived from the supply and demand functions, the system obtained by substituting the supply and demand functions into this set of adding-up restrictions exhausts the equilibrium conditions of the model:

$$\sum_y \sum_h \exp\left(m_{\theta_{xyh}}(1-\epsilon_{\theta_{xyh}})\Phi_{xyh}\right) a_x = 1 \quad \text{for all } x, \tag{3.18}$$

$$\sum_x \sum_h \exp\left(q_{\theta_{xyh}}\epsilon_{\theta_{xyh}}\Phi_{xyh}\right) b_y = 1 \quad \text{for all } y, \tag{3.19}$$

$$\text{where: } \theta_{xyh} = \exp\left([q_{\theta_{xyh}}\epsilon_{\theta_{xyh}} - m(\theta_{xyh})(1-\epsilon(\theta_{xyh}))]\Phi_{xyh}\right)\frac{f_y}{n_x}\frac{b_y}{a_x}. \tag{3.20}$$

Equilibrium: Existence and Uniqueness. We show existence of a unique equilibrium in appendix A.3. We do so by first showing that that a unique solution θ^* exists that satisfies the market tightness equation (3.20). This solution can then be plugged into the mappings (3.18) and (3.19), which we show to be a contracting system of $X+Y$ equations in the same number of unknowns a_x and b_y. This also highlights the approach we use to compute the equilibrium. We start from an arbitrary guess on the parameters and on market tightness. For these values, we can solve the system defined by (3.18) and (3.19) to obtain a solution for a_x and b_y. This solution allows us to update the value for market tightness through (3.20). These steps are iterated until convergence.

4 The Data

4.1 Data

Labor Market and Unemployment. We use the latest wave of data (2022) from the American Community Survey (ACS) provided by the Integrated Public Use Microdata Series (IPUMS). The ACS is representative for the United States in terms of demographics and labor market outcomes. We select a sample of individuals in the labor force who are between the ages of 25 and 64. We remove individuals that are self-employed or in marginal employment – defined as working less than ten hours per week or less than forty weeks per year. We only keep individuals who hold a college degree and are currently, or were in the last five years, employed in a high earning occupation as defined in Goldin (2014). This yields a sample of 130,934 respondents.

Job Vacancies. We construct the occupational distribution of job vacancies within each state based on job postings from the National Labor Exchange (NLx) Research Hub dataset. The NLx data collects vacancies from companies' career websites, states' job vacancy banks, and the United States federal jobs portal, and attempts to cover all real job postings in the United States. Roughly one third to one half of the job vacancies in the Job Openings and Labor Turnover Survey (JOLTS) are accounted for in NLx openings[8]. We scale the vacancy distributions through aggregate counts for each state obtained from JOLTS. More information on this data can be found in Appendix A.4.1.

Aggregation into Types. We aggregate workers, hours, and occupations into discrete types. Workers are characterized by their gender, potential experience, and type of degree. We construct potential experience by discretizing the measure of Acemoglu and Autor (2011) into more, or less than ten years.[9] The type of degree distinguishes between low and high earning fields. This distinction is made using k-means clustering of average earnings for each degree, similar to Almar et al. (2023). This leads to $X = 8$ different types of workers.[10] We discretize hours as in Bick et al. (2022), who distinguish between part-time (less than 40 hours per week), full-time (between 40 and 50 hours per week), and a long hour (more than 50 hours per week) positions,

[8] The numbers differ between states and increased substantially over the years since the early 2010s.

[9] Their measure is an updated version of Park (1994), which relies on information about average degree completion times across demographic groups to construct potential experience.

[10] The definition and number of types is chosen to maximize the explained wages, while keeping the model tractable for our structural estimation.

such that $H = 3$. The occupations are defined by the categories introduced in Goldin (2014), which distinguish between business, health, technology, science, and other occupations, such that we have $Y = 5$ types. This means that we have 120 sub-markets. We observe these submarkets across different independent markets z, which are distinguished by their census divisions (aggregating several states) of which there are $Z = 9$.

4.2 Summary Statistics

Table 1 presents a summary of the estimation sample, which consists of almost 135,000 workers in the United States. The vast majority of them are employed ($\approx$ 98.5%) because the sample consists of higher educated workers who (used to) work in high earning occupations. Because we focus on higher educated workers, women slightly outnumber men, at 52.7 as opposed to 47.3%. But men are significantly more likely to have a high earning degrees than women, at 51% as opposed to 21%, while experience levels are relatively similar. The main difference in terms of occupations is that women are substantially more likely to work in health occupations than men, who work more in tech and other occupations. When we look at working hours, we find that almost no men work part-time, compared to roughly 10% of women. On the other hand, only 16% of women work part-time, as opposed to 27% of men. The raw gender gap in monthly wages is about 40%, which shrinks to roughly 30% when we look at hourly wages.

We disaggregate the data further in Figure 1. Panel a highlights differences in long hours across worker types and occupations. The first takeaway is that hours differ substantially across occupations. This is the focus of recent work by Erosa et al. (2022). We find that in Business and Other occupations – which notably includes law occupations – about a quarter of employees works long hours.[11] In Tech and Science there are already substantially less employees working long hours (about 20%). In Health occupations the fraction that works long hours is only 10%.

The second main takeaway is that hours differ substantially *within* occupations across worker types. For example, in Business, more than one in three men (M) with a high earning degree (HD) and high experience (HE) work long hours. This is more than double the fraction of women (W) in the same group, among whom 16% works long hours, and of men with a low earning degree

[11] The group of 'Other' occupations contains law occupations, but also construction managers, non-retail sales workers, licensing inspectors, and architects.

Table 1: Summary Statistics

	Men	Women	Total
Observations			
Employed Respondents	61,921	69,013	130,934
Unemployed Respondents	884	967	1,851
Demographics			
High Earn. Degree	0.51	0.21	0.35
High Exp.	0.78	0.76	0.77
Occupations			
Business	0.49	0.44	0.47
Health	0.11	0.4	0.26
Tech	0.22	0.06	0.14
Science	0.04	0.03	0.03
Other	0.14	0.08	0.11
Hours			
Part-Time	0.02	0.1	0.06
Full-Time	0.7	0.74	0.72
Long Hours	0.27	0.16	0.22
Wages			
Avg. Month. Wage (in k$)	6.49	4.6	5.49
Std. Dev. Month. Wage (in k$)	4.72	3.08	4.06
Avg. Hr. Wage (in $)	34.13	26.25	29.98
Std. Dev. Hr. Wage (in $)	22.4	15.32	19.41

Notes. Summary statistics for the main estimation of workers from the United States. Based on data from the American Community Survey (ACS, 2022). The selected sample consists of higher educated workers in specific occupations as in Goldin (2014) – see the discussion in section 4.1.

(LD) and low experience (LE), among whom 23% works long hours. While the differences are the largest in Business and Other occupations, we also find large differences in health, where about 7% of women with low experience and a low earning degree work long hours, as opposed to men with a high earning degree and low experience, where the fraction is just over 20%. We find comparable differences in Tech, but not in Science occupations, where differences across worker types are relatively small. Panel b show that part-time work also differs substantially between occupations and across worker types. In all occupations, except for Health, the fraction working part-time is relatively small. But even within Health occupations, essentially only women work part-time.

Looking at earnings in Panel c we again see substantial heterogeneity, with Business and

Figure 1: Hours and Wages by Types

(a) Long Hours by Type

	Business	Health	Tech	Science	Other
F-LD-HE	0.21	0.12	0.18	0.16	0.2
M-LD-HE	0.32	0.19	0.23	0.17	0.31
F-LD-LE	0.16	0.07	0.12	0.14	0.15
M-LD-LE	0.23	0.12	0.16	0.2	0.24
F-HD-HE	0.24	0.12	0.17	0.18	0.22
M-HD-HE	0.36	0.2	0.23	0.19	0.36
F-HD-LE	0.21	0.12	0.12	0.17	0.2
M-HD-LE	0.29	0.21	0.16	0.2	0.3
Average	0.26	0.12	0.2	0.18	0.27

(b) Part-Time by Type

	Business	Health	Tech	Science	Other
F-LD-HE	0.06	0.18	0.03	0.08	0.05
M-LD-HE	0.02	0.06	0.01	0.03	0.02
F-LD-LE	0.04	0.14	0.02	0.06	0.05
M-LD-LE	0.01	0.05	0.03	0.02	0.03
F-HD-HE	0.06	0.18	0.05	0.04	0.04
M-HD-HE	0.02	0.06	0.01	0.02	0.01
F-HD-LE	0.02	0.1	0.02	0.06	0.04
M-HD-LE	0.02	0.06	0.01	0.09	0.01
Average	0.03	0.15	0.02	0.05	0.03

(c) Earnings by Type

	Business	Health	Tech	Science	Other
F-LD-HE	4.95	4.23	5.36	4.71	4.56
M-LD-HE	6.86	5.64	6.36	5.25	5.56
F-LD-LE	3.6	3.51	3.92	3.02	3.25
M-LD-LE	4.05	4.2	4	3.14	3.51
F-HD-HE	6.03	5.09	6.67	5.68	5.8
M-HD-HE	8.58	6.66	7.36	6.8	6.81
F-HD-LE	4.46	4.27	4.47	3.54	3.85
M-HD-LE	5.19	4.54	4.57	3.79	4.31
Average	5.95	4.41	6.34	4.9	5.22

Notes. Hours and wages by worker and occupation type in the United States. Data from the American Community Survey (2022). Panel (a) shows the distribution of employees that work long hours (> 50 per week) across different types of workers and occupations. Panel (b) shows the same distribution but for part-time hours (< 35 per week). Panel (c) shows the distribution of monthly wages (in k$) across the same set of types. The abbreviations in worker types distinguish between men and women (M/F), low and high earning degrees (LD/HD), and low and high experience (LE/HE).

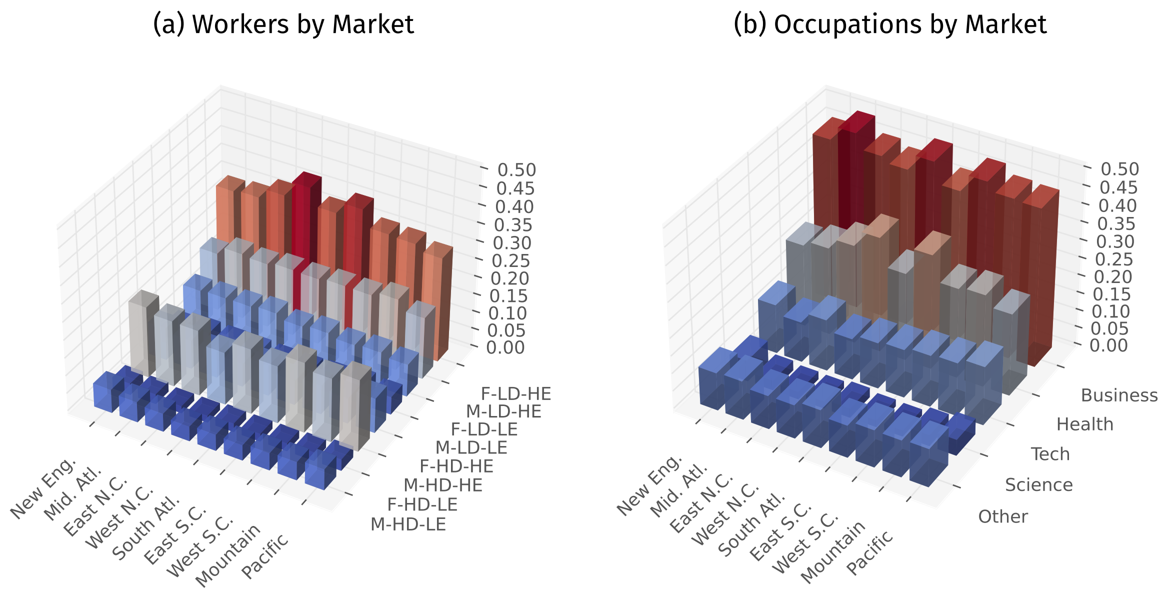
Tech occupations paying the highest monthly wages (on average roughly 6,000$). The average across the other occupations is roughly 1,000$ lower. We see that men with high earning degrees and high experience earn the highest salaries on average, and particularly in Business and Tech occupations. Women and men with low earning degrees and low experience earn the lowest average salaries, with no notable differences across occupations. There are again substantial differences between groups within each occupations, with the largest differences again being in Business occupations, where men with high earning degrees and high experience earn more than twice as much (about 8,600$) than men or women with low earning degrees and low experience (between 3,600 and 4,000$). This difference is notable because it does not reflect differences in part-time work, which is low in all these groups as seen in Panel b.

Heterogeneity Across Markets. We plot the marginal distributions – the sum of matched and unmatched workers and firms – across occupations and worker types for different markets in Figure 2. We need some variation in types across markets to identify the search friction parameter, as discussed below in section 5. When the marginals are identical, the same primitives should lead to the same outcomes in terms of matching patterns and wages, and we would have no variation to exploit. We find that the distribution of both worker and firm types are overall quite similar across markets but with some notable heterogeneity. For example, in Panel a we see that in the West North Central market, roughly 35% of the population consists of women with low earning degrees but high experience. In the Pacific market this group makes up only about 25% of the population. We find differences in the distribution of occupations of similar order in Panel b. In the Mid Atlantic market about 47% of employees in our sample are employed in Business occupations, whereas in the East South Central market this is only about 40%.

5 Maximum Likelihood Estimation

Observations. We observe individual-level data on matches between workers and firms in several independent markets $z \in Z$. This allows us to recover matching patterns by type, denoted by $\hat{\mathcal{M}}_{xyhz}$ for (x, y, h) matches, $\hat{\mathcal{M}}_{x0z}$ for unemployed workers of type x, and $\hat{\mathcal{M}}_{0yz}$ for vacancies of type y. We denote the marginals by $\mathcal{N}_{xz}$ and $\mathcal{F}_{yz}$. We also observe a noisy measure of wages

Figure 2: Workers and Firms across Markets



Notes. This figure documents the total number of workers (both employed and unemployed) and jobs (both filled and vacant) of each type across Census Divisions in the United States. The data used to construct this figure are a combination of the American Community Survey (2022), the Job Openings and Labor Turnover Survey, and the job vacancy database by the National Labor Exchange's Research Hub. The abbreviations in worker types distinguish between men and women (M/F), low and high earning degrees (LD/HD), and low and high experience (LE/HE). These are the values referred to in the model by $\mathcal{N}_{xz}$ and $\mathcal{F}_{yz}$.

$\hat{\mathcal{W}}_{ijkz}$ that relates to the true equilibrium wage by:

$$\hat{\mathcal{W}}_{ijkz} = w_{xyhz} + \delta_{ijkz}, \text{ with } \delta_{ijkz} \sim \mathcal{N}(0, s) \text{ iid}, \tag{5.1}$$

where δ_{ijkz} is a centered Gaussian measurement error with standard deviation s.[12]

Identification. We briefly discuss how the model is identified with our data from several identical (in terms of parameters) but independent markets z. Let Ω denote the complete set of parameters in the model, which are workers' amenity values (α_{xyh}), firms' productivity values (ρ_{xyh}), and the elasticity parameter of the matching function (λ). First note that both labor supply and demand increase in the joint matching surplus Φ_{xyh}. The equilibrium matching patterns thus inform the model about the match surplus as in Choo and Siow (2006). In addition to matching patterns, we also observe equilibrium wages. These allow us to separate supply-

[12]We can allow for a richer measurement error distribution, for example by allowing for the error variance to depend on (x, y, h).

from demand side valuations as in Dupuy and Galichon (2022). To see this, note that equilibrium wages in equation (3.5) increase in firms' productivity value ρ_{xyh} but decrease in workers' amenity value α_{xyh}. The final parameter we need to identify is the matching function elasticity λ. This parameter governs the shape of the worker-firm bargaining frontier – we can see this explicitly by substituting (3.8) into (3.9). As discussed in Galichon et al. (2019) variation across independent markets can be exploited to matching patterns across independent markets to identify this parameter. More details can be found in Appendix A.3.1.

Parameterization. We parameterize the model with a flexible specification. The value of amenities is modeled as:

$$\alpha_{xyh} = \sum_{x,y,h} \mathbb{1}_{\alpha_{x \times y \times h}}, \tag{5.2}$$

where $\mathbb{1}_{\alpha_{x \times y \times h}}$ are dummy variables for each worker - occupation - hour combination. The disutility of working a given number of hours is thus specific to each worker-firm match.

The productivity function is parameterized as:

$$\rho_{xyh} = \sum_{x} \mathbb{1}_{\rho_x} + \sum_{y,h} \mathbb{1}_{\rho_{y \times h}}, \tag{5.3}$$

where $\mathbb{1}_{\rho_x}$ are dummy variables for the different types of workers. We thus allow for heterogeneity in how firms value workers of different experience and with different degrees. The second set of dummies $\mathbb{1}_{\rho_{y \times h}}$ are for occupation - hour combinations, allowing for the productivity of hours to be occupation-specific.

5.1 Estimation

We estimate the parameters of the model through maximum likelihood. The estimation procedure consists of solving the model for a given parameter vector Ω and using the model equations to equate the predicted matching patterns and wages to the observed ones.

The Likelihood Function. We split the likelihood function into three components. The likelihood contribution of employed individuals consists of a part that aims to equate the predicted matching distribution to its sample counterpart and a component that aims to equate predicted to observed wages. The contribution of unemployed individuals only aims to equate predicted and observed matching patterns. Note that we need to integrate out over their search decisions, on which we have no data. The contributions of vacant positions are analogous.

The log likelihood contribution of an individual observation can thus be expressed as for each market z as:

$$\begin{aligned}\log l_{ijk}(\Omega) = \mathbb{1}_{ijk=xyh} &\left(\log M_{xyh|\Omega} - \frac{\left(\hat{w}_{ijk} - W_{xyh|\Omega}\right)^2}{2s^2} - \log s \right) \\ &+ \mathbb{1}_{i=x0} \left(\log \sum_{y,h} M_{x0|\Omega} \right) \\ &+ \mathbb{1}_{j=0y} \left(\log \sum_{x,h} M_{0y|\Omega} \right),\end{aligned} \tag{5.4}$$

where $\mathbb{1}_{ijk=xyh}$ is an indicator that equals one if observation ijk is a match of type (x, y, h), $\mathbb{1}_{i=x0}$ is an indicator for observation i being an unemployed person of type x, and $\mathbb{1}_{j=0y}$ is an indicator for observation j being a vacant position of type y. Taking account of our sample consisting of several independent markets z, the estimated parameter vector $\hat{\Omega}$ solves:

$$\hat{\Omega} = \arg\max_{\Omega} \left\{ \sum_z \sum_i \sum_j \sum_k \log l_{ijk}(\Omega) \right\}. \tag{5.5}$$

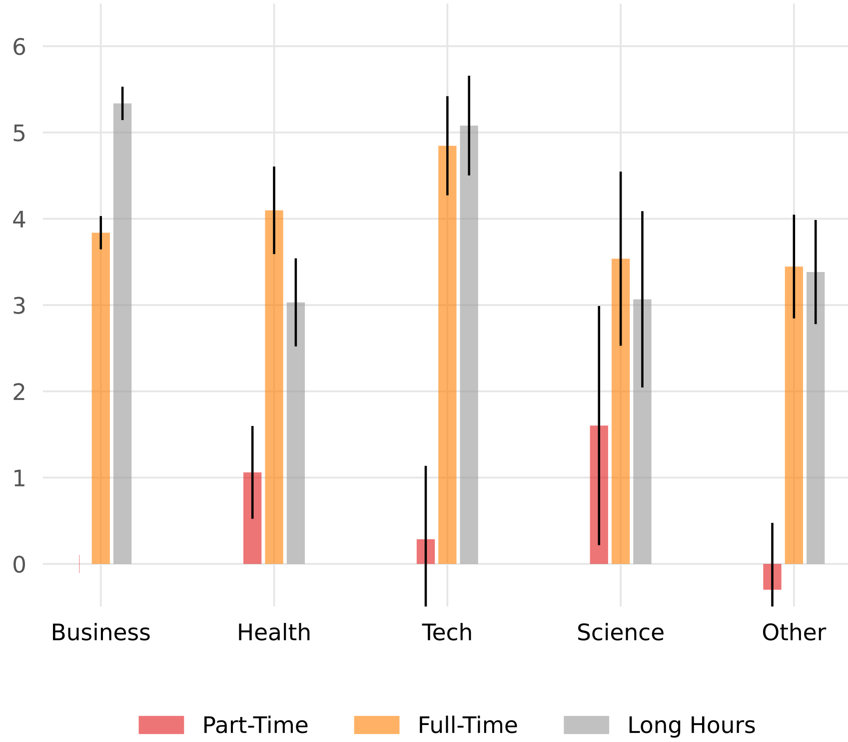
6 Results

6.1 Model Estimates

Productivity Parameters. We first look at how the (perceived) productivity of hours differs across occupations in Figure 3. The first main result is consistent with the narrative of Goldin (2014). We find that only in Business occupations long hours are substantially more productive than standard full time hours. Here employees generate on average about 55.000$ per month when working long hours compared to less than 40.000$ when working full time. As discussed earlier this is reflected in substantial wage premiums (see section 4.1). In all other occupations, differences in the perceived productivity of full-time and longer hours are negligible.

A second pattern in Figure 3 that is also consistent with Goldin (2014) is that part-time work

Figure 3: Productivity across Hours and Occupations

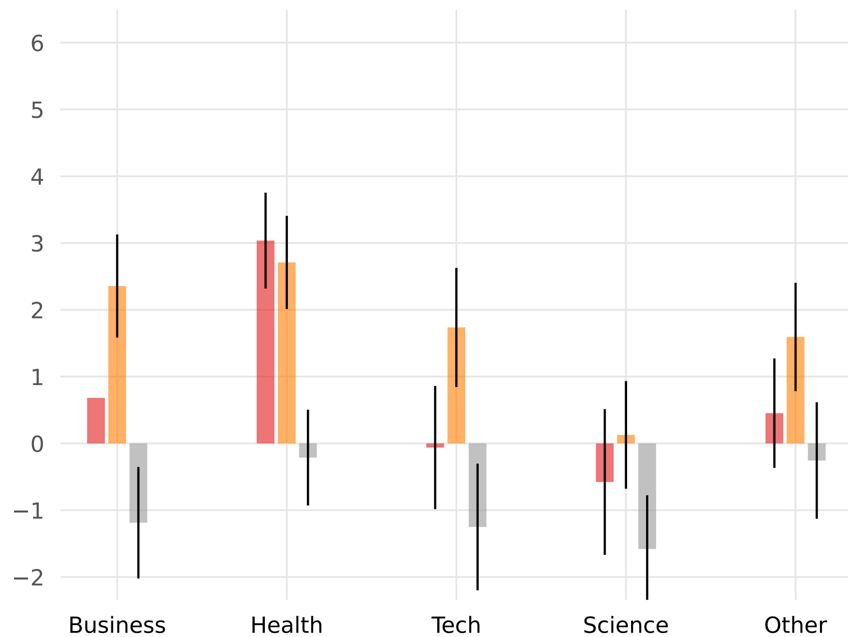


Notes. This figure shows the productivity estimates of hours for each occupation group. These are the estimates of $\rho_{y\times h}$ in equation (5.2). Part-time employment in Business occupations is the omitted reference group.

is the *least* productive in Business and Other (which includes law) occupations. We find that in Health and Science occupations, part-time work is substantially more productive than in Business, and long hours are the least productive relative to full-time work. A notable group within the health occupations are pharmacists, which is discussed in Goldin and Katz (2016) as being characterized by relatively small productivity premiums due to low information costs associated with the extensive use of computer systems to track clients.

Amenity Valuations. We now look at the final determinant of occupational hours and wages in the model – how workers value hours differently across occupations. We interpret occupational heterogeneity in the perceived amenity value as reflecting, for example, differences in other unobserved aspects of temporal (e.g. being on call) or location (e.g. working from home) flexibility that affect how workers feel about the number of hours. Figure 4 shows the average amenity values across occupations and hours, integrating out over the worker type distribution. The first main takeaway is that worker-side differences are about half as large in magnitude as firm-side heterogeneity in productivity. As expected, we find that workers dislike long hours more than working full-time, particularly in Business, Tech, and Science occupations. We also find a strong preference for full-time over part-time hours, which suggests that workers are willing to forego some hourly wages for the increase in earnings associated with working more hours – this is consistent with the literature, see for example Maestas et al. (2023). We find the largest effects in health occupations, where workers in part-time and full-time positions perceive their amenities as particularly valuable. One interpretation for this result is that they perceive their work as particularly meaningful – see Cassar and Meier (2018).

Figure 4: Amenities across Hours and Occupations



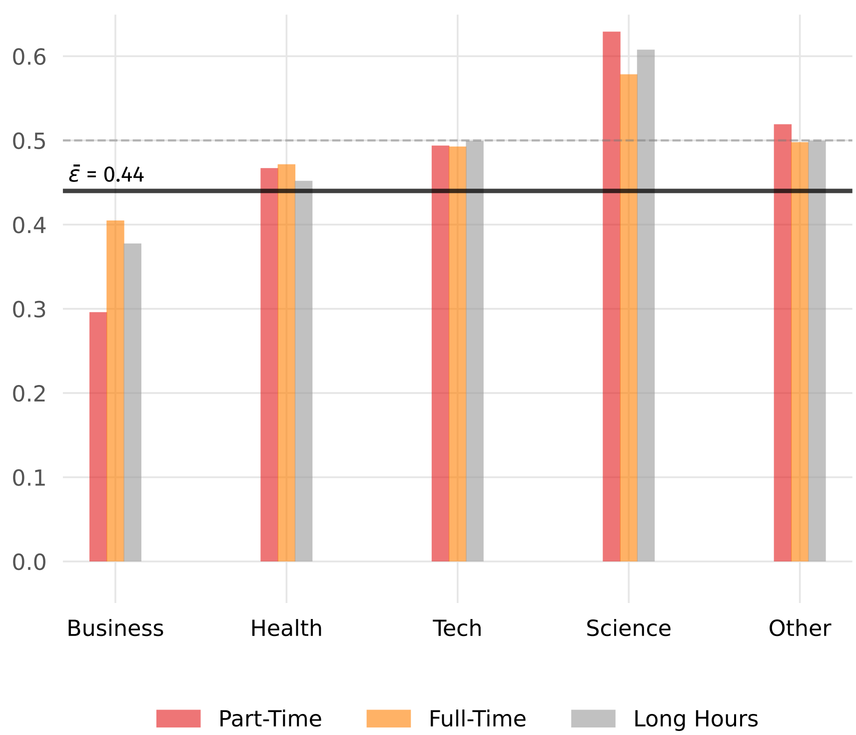
Notes. This figure shows the productivity estimates of hours for each occupation group. These are the estimates of $\alpha_{y\times h}$ in equation (5.2). Part-time employed men with low earning degrees and low experience in Business occupations the omitted as a reference group.

Market Structure and Bargaining Weights. The final determinant of hours and wages in the model are the market structure effects that determine the bargaining weights. We now look at how the weights differ across occupations and hours by integrating out over the different markets and workers:

$$\epsilon_{yh} = \sum_{x,z} \frac{\epsilon_{xyhz} \times M_{xyhz}}{n_{xz}}. \tag{6.1}$$

We plot the bargaining weights in Figure 5. Note that these measures the firm's bargaining power, with firms being relatively more powerful if $\epsilon_{yh} > 0.5$, and lie on the unit interval. First of all, note that the average aggregate weight $\bar{\epsilon}$ that we obtain by also integrating out occupations and hours is 0.44. This reflects the relatively tight nature of the labor market in the United States (see for example Autor et al. (2023)). But the figure reveals substantial occupational heterogeneity. In Business occupations firms' bargaining weights only lie between 0.3 and 0.4. The excess demand in these occupations thus further pushes up the wages. At the other end we have Science occupations, where firms' bargaining weights are roughly 0.6, pushing down wages. In Health, Tech, and Other occupations bargaining power is relatively evenly distributed.

Figure 5: Bargaining Weights across Hours and Occupations



Notes. This figure shows the bargaining weight estimates $\epsilon_{\theta_{yh}}$ by hours and occupation group. We average across worker types and markets through equation (6.1).

6.2 Model Fit

We now look at Figure 6 to study how the model's predicted wages and matching patterns compare to those in the sample. The diagonal lines indicate a perfect model fit. We find that the model replicates matching patterns in the data very well. The number of workers across hours and occupations (Panel a) are very well replicated. The number of job vacancies across occupations also matches well but the model generally predicts too many vacancies (Panel d). The model also over predicts the degree of unemployment for most types of workers (Panel c). When we look at the wage distribution we find that the model generally does not generate enough dispersion in wages (Panel b), and predicts a wage distribution that is too compressed.

6.3 Counterfactual Experiments

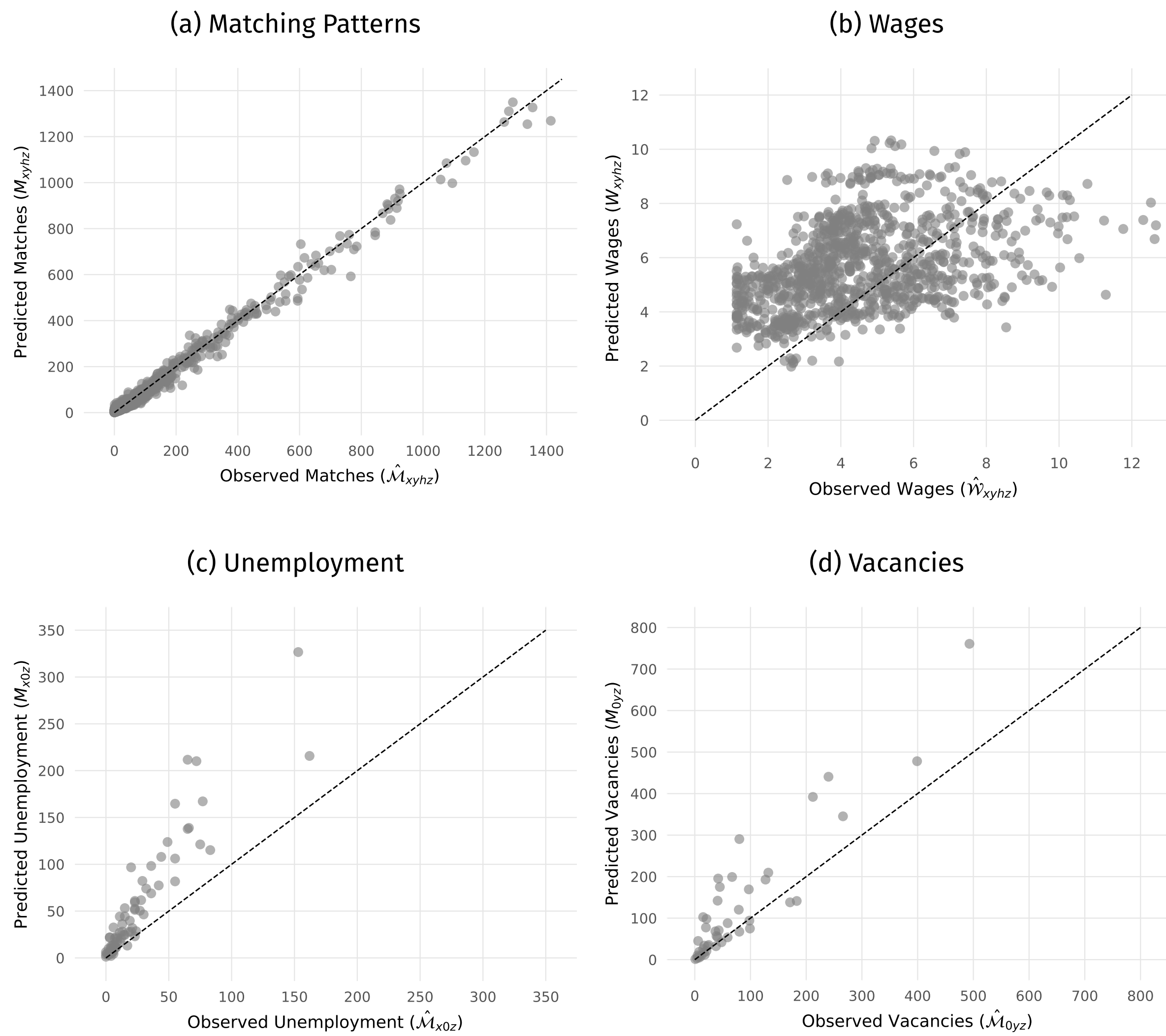
We consider two counterfactual scenarios. The first studies what happens when we change the productivity of long hours in Business occupations. This allows us to test Goldin (2014)'s hypothesis that the gender gap in earnings would be considerably reduced – and may disappear altogether – if the incentives to reward long hours were not so high in Business occupations. This counterfactual can be feasible because similar changes have occurred in other occupations as shown in Goldin and Katz (2016). We operationalize it by setting the productivity of long hours (ρ_{yh}) in Business occupations equal to that of full time hours. The second counterfactual studies the impact of bargaining power and considers what happens when we would equalize the bargaining weights of workers and firms across all occupations. [to finish]

6.4 What Explains Productivity and Premiums and Penalties?

We now study which occupational traits can predict productivity premiums to long hours and penalties to part-time work. To study this we regress a large set of job characteristics (skills, work activities, and work styles) from the ONET database on the estimated parameters:

$$\rho_{yh} = c + \alpha Skills_{ij} + \beta WorkActivities_{ij} + \gamma WorkStyles_{ij} + \epsilon_{ij}. \tag{6.2}$$

Figure 6: Matching Patterns and Wages (observed vs. predicted)



Notes. Model predictions and data moments on matching patterns and wages. The scatters are moments by worker type $\times$ firm type $\times$ hours type $\times$ market, with dots on the 45° line indicating a perfect model replication of the empirical patterns. The matching patterns are expressed in levels and monthly wages in 1000s of dollars.

We select a model using the LASSO procedure outlined in Belloni et al. (2012), which allows for heteroskedastic errors and uses a data-driven method to select the penalty. [to finish]

7 Conclusion

This paper introduces a tractable search and matching model with endogeneous equilibrium hours and wages. We build on the separable extreme value approach introduced in Choo and

Siow (2006) but introduce search frictions as in Moen (1997). We use the model to study differences in productivity, amenities, and bargaining power across occupations. The estimates reveal substantial heterogeneity in the productivity of hours across occupations, and the main patterns are in line with the hypothesis put forth in Goldin (2014). We also find an important role for bargaining power, further pushing up wages in occupations where demand exceeds supply.

To conclude, we point out some directions for future research. A first extension would make the model dynamic. This allows us to take into account how different hours choices can affect human capital accumulation over the life cycle as in Michelacci and Pijoan-Mas (2012). Predicted future benefits may be an important reason for workers to choose long hours. A second extension of the model can introduce a model of the household. This opens up interesting questions about the interaction between intra-household and labor market bargaining.

References

Aaronson, Daniel and Eric French (2004) "The effect of part-time work on wages: Evidence from the social security rules," *Journal of Labor Economics*, 22 (2), 329–352.

Acemoglu, Daron and David Autor (2011) "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of labor economics*, 4 (b), 1044–1171.

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo (2022) "Artificial intelligence and jobs: Evidence from online vacancies," *Journal of Labor Economics*, 40 (S1), S293–S340.

Almar, Frederik, Benjamin Friedrich, Ana Reynoso, Bastian Schulz, and Rune Majlund Vejlin (2023) "Marital sorting and inequality: How educational categorization matters,"Technical report, IZA Discussion Papers.

Arcidiacono, Peter, Andrew Beauchamp, and Marjorie McElroy (2016) "Terms of endearment: An equilibrium model of sex and matching," *Quantitative Economics*, 7 (1), 117–156.

Autor, David, Arindrajit Dube, and Annie McGrew (2023) "The unexpected compression: Competition at work in the low wage labor market,"Technical report, National Bureau of Economic Research.

Autor, David H and David Dorn (2013) "The growth of low-skill service jobs and the polarization of the US labor market," *American economic review*, 103 (5), 1553–1597.

Beauchamp, Andrew, Rossella Calvi, and Scott Fulford (2024) "Terms of Engagement: Migration, Dowry, and Love in Indian Marriages."

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen (2012) "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 80 (6), 2369–2429.

Bick, Alexander, Adam Blandin, and Richard Rogerson (2022) "Hours and wages," *The Quarterly Journal of Economics*, 137 (3), 1901–1962.

Cassar, Lea and Stephan Meier (2018) "Nonmonetary incentives and the implications of work as a source of meaning," *Journal of Economic Perspectives*, 32 (3), 215–238.

Chade, Hector, Jan Eeckhout, and Lones Smith (2017) "Sorting through search and matching models in economics," *Journal of Economic Literature*, 55 (2), 493–544.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir (2020) "Changes in assortative matching: Theory and evidence for the US,"Technical report, National Bureau of Economic Research.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss (2017) "Partner choice, investment in children, and the marital college premium," *American Economic Review*, 107 (8), 2109–2167.

Choo, Eugene and Aloysius Siow (2006) "Who marries whom and why," *Journal of political Economy*, 114 (1), 175–201.

Ciscato, Edoardo (2024) "Assessing racial and educational segmentation in large marriage markets."

Cogan, John (1981) "Fixed costs and labor supply," *Econometrica*, 49 (4), 945–963.

Corblet, Pauline (2023) "Education Expansion, Sorting, and the Decreasing Wage Premium."

Cortés, Patricia and Jessica Pan (2019) "When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap," *Journal of Labor Economics*, 37 (2), 351–398.

Deming, David J and Kadeem Noray (2020) "Earnings dynamics, changing job skills, and STEM careers," *The Quarterly Journal of Economics*, 135 (4), 1965–2005.

Dupuy, Arnaud (2021) "Migration in China: To work or to wed?" *Journal of Applied Econometrics*, 36 (4), 393–415.

Dupuy, Arnaud and Alfred Galichon (2014) "Personality traits and the marriage market," *Journal of Political Economy*, 122 (6), 1271–1319.

——— (2022) "A note on the estimation of job amenities and labor productivity," *Quantitative Economics*, 13 (1), 153–177.

Dupuy, Arnaud, Alfred Galichon, Sonia Jaffe, and Scott Duke Kominers (2020) "Taxation in matching markets," *International Economic Review*, 61 (4), 1591–1634.

Dupuy, Arnaud, John Kennes, and Ran Sun Lyng (2023) "Job Amenities in the Market for CEOs."

Eeckhout, Jan and Philipp Kircher (2010) "Sorting and decentralized price competition," *Econometrica*, 78 (2), 539–574.

Ekeland, Ivar, James J Heckman, and Lars Nesheim (2004) "Identification and estimation of hedonic models," *Journal of political economy*, 112 (S1), S60–S109.

Erosa, Andrés, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson (2022) "Hours, occupations, and gender differences in labor market outcomes," *American Economic Journal: Macroeconomics*, 14 (3), 543–590.

——— (2024) "Wage and earnings inequality between and within occupations: The role of labor supply," *Journal of Monetary Economics*, 103553.

French, Eric (2005) "The effects of health, wealth, and wages on labour supply and retirement behaviour," *The Review of Economic Studies*, 72 (2), 395–427.

Galichon, Alfred, Scott Duke Kominers, and Simon Weber (2019) "Costly concessions: An empirical framework for matching with imperfectly transferable utility," *Journal of Political Economy*, 127 (6), 2875–2925.

Galichon, Alfred and Bernard Salanié (2022) "Cupid's invisible hand: Social surplus and identification in matching models," *The Review of Economic Studies*, 89 (5), 2600–2629.

Goldin, Claudia (2014) "A grand gender convergence: Its last chapter," *American economic review*, 104 (4), 1091–1119.

Goldin, Claudia and Lawrence F Katz (2016) "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation," *Journal of Labor Economics*, 34 (3), 705–746.

Hershbein, Brad and Lisa B Kahn (2018) "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings," *American Economic Review*, 108 (7), 1737–1772.

Jaffe, Sonia and Simon Weber (2019) "The effect of meeting rates on matching outcomes," *Economic Theory*, 67 (2), 363–378.

Jang, Youngsoo and Minchul Yum (2022) "Nonlinear occupations and female labor supply over time," *Review of Economic Dynamics*, 46, 51–73.

Kaas, Leo and Philipp Kircher (2015) "Efficient firm dynamics in a frictional labor market," *American Economic Review*, 105 (10), 3030–3060.

Kuhn, Peter and Fernando Lozano (2008) "The expanding workweek? Understanding trends in long work hours among US men, 1979–2006," *Journal of Labor Economics*, 26 (2), 311–343.

Legros, Patrick and Andrew F Newman (2007) "Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities," *Econometrica*, 75 (4), 1073–1102.

Macpherson, David A and Barry T Hirsch (1995) "Wages and gender composition: why do women's jobs pay less?" *Journal of labor Economics*, 13 (3), 426–471.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger (2023) "The value of working conditions in the United States and implications for the structure of wages," *American Economic Review*, 113 (7), 2007–2047.

Manning, Alan and Barbara Petrongolo (2008) "The part-time pay penalty for women in Britain," *The economic journal*, 118 (526), F28–F51.

Mantovani, Cristiano (2023) "Hours-Biased Technological Change,"Technical report, Working Paper.

Menzio, Guido and Shouyong Shi (2010) "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 145 (4), 1453–1494.

Michelacci, Claudio and Josep Pijoan-Mas (2012) "Intertemporal labour supply with search frictions," *The Review of Economic Studies*, 79 (3), 899–931.

Moen, Espen R (1997) "Competitive search equilibrium," *Journal of political Economy*, 105 (2), 385–411.

Mourifié, Ismael and Aloysius Siow (2021) "The Cobb-Douglas marriage matching function: Marriage matching with peer and scale effects," *Journal of Labor Economics*, 39 (S1), S239–S274.

Park, Jin Huem (1994) "Estimation of Sheepskin Effects and Returns to Schooling Using he Old and the New CPS Measures of Educational Attainment,"Technical report.

Rosen, Sherwin (1974) "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal of political economy*, 82 (1), 34–55.

——— (1986) "The theory of equalizing differences," *Handbook of labor economics*, 1, 641–692.

Roth, Alvin E and Marilda Sotomayor (1989) "The college admissions problem revisited," *Econometrica*, 559–570.

Roy, Andrew Donald (1951) "Some thoughts on the distribution of earnings," *Oxford economic papers*, 3 (2), 135–146.

San, Shmuel (2024) "Who Works Where and Why: The Role of Social Connections in the Labor Market."

Schaal, Edouard (2017) "Uncertainty and unemployment," *Econometrica*, 85 (6), 1675–1721.

Shimer, Robert and Lones Smith (2000) "Assortative matching and search," *Econometrica*, 68 (2), 343–369.

Siow, Aloysius (2015) "Testing Becker's theory of positive assortative matching," *Journal of Labor Economics*, 33 (2), 409–441.

Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri (2021) "Directed search and competitive search equilibrium: A guided tour," *Journal of Economic Literature*, 59 (1), 90–148.

Yurdagul, Emircan (2017) "Production complementarities and flexibility in a model of entrepreneurship," *Journal of Monetary Economics*, 86, 36–51.

A Appendix

A.1 Equilibrium Solution

We start from the equilibrium optimization problem:

$$\max_{\theta,w,x,h}\left\{\frac{m(\theta_{xyh})}{\theta_{xyh}}(\rho_{xyh}-W_{xyh})+\eta_{xy_jh}\right\}$$
$$\text{s.t. }\; m(\theta_{xyh})(W_{xyh}+\alpha_{xyh})+\xi_{x_iyh}=u_{x_iyh}$$

We can substitute out the constraint to obtain:

$$\max_{\theta,x,h}\left\{\frac{m(\theta_{xyh})(\rho_{xyh}+\alpha_{xyh})-u_{x_iyh}+\xi_{x_iyh}+\eta_{xy_jh}\theta_{xyh}}{\theta_{xyh}}\right\} \tag{A.1}$$

The first order condition with respect to market tightness is:

$$u_{x_iyh}=\Phi_{xyh}\left(m(\theta_{xyh})-\theta_{xyh}m'(\theta_{xyh})\right)+\xi_{x_iyh} \tag{A.2}$$

We can substitute this into the constraint ot obtain a characterization of equilibrium wages:

$$W_{xyh}=\left(1-\epsilon(\theta_{xyh})\right)\rho_{xyh}-\epsilon(\theta_{xyh})\nu_{xyh}, \tag{A.3}$$

where $\epsilon(\theta_{xyh})$ is defined by:

$$\epsilon(\theta_{xyh}):=\theta_{xyh}\frac{m'(\theta_{xyh})}{m(\theta_{xyh})}. \tag{A.4}$$

A.2 Equilibrium Solution with Taxation

We now introduce proportional income taxation into the model as in Dupuy et al. (2020):

$$\max_{\theta,w,x,h} \left\{ \frac{m(\theta_{xyh})}{\theta_{xyh}}(\rho_{xyh} - W_{xyh}) + \eta_{xy_jh} \right\}$$
$$\text{s.t. } m(\theta_{xyh})(W_{xyh} + (1-\tau)\alpha_{xyh}) + \xi_{x_iyh} = u_{x_iyh}$$

We can substitute out the constraint to obtain:

$$\max_{\theta,x,h} \left\{ \frac{m(\theta_{xyh})((1-\tau)\rho_{xyh} + \alpha_{xyh}) - u_{x_iyh} + \xi_{x_iyh} + \eta_{xy_jh}\theta_{xyh}}{\theta_{xyh}} \right\} \tag{A.5}$$

The first order condition with respect to market tightness is:

$$u_{x_iyh} = \Phi_{xyh}\left(m(\theta_{xyh}) - \theta_{xyh}m'(\theta_{xyh})\right) + \xi_{x_iyh} \tag{A.6}$$

We can substitute this into the constraint ot obtain a characterization of equilibrium wages:

$$W_{xyh} = (1 - \epsilon(\theta_{xyh}))\,\rho_{xyh} - \epsilon(\theta_{xyh})\frac{\alpha_{xyh}}{1-\tau}, \tag{A.7}$$

where $\epsilon(\theta_{xyh})$ is defined by:

$$\epsilon(\theta_{xyh}) := \theta_{xyh}\frac{m'(\theta_{xyh})}{m(\theta_{xyh})}. \tag{A.8}$$

A.3 Equilibrium Existence and Uniqueness

Consider the following system of equations that exhaust the equilibrium conditions of the model:

$$\sum_y \sum_h \exp\left((1+\theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda}\Phi_{xyh}\right) a_x = 1 \quad \text{for all } x, \tag{A.9}$$

$$\sum_x \sum_h \exp\left(\theta_{xyh}^{\lambda-1}(1+\theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda}\Phi_{xyh}\right) b_y = 1 \quad \text{for all } y, \tag{A.10}$$

$$\text{where: } \theta_{xyh} = \exp\left(\frac{\theta_{xyh}^{\lambda-1}-1}{(1+\theta_{xyh}^{\lambda})^{(\lambda-1)/\lambda}}\Phi_{xyh}\right)\frac{a_x}{b_y}\frac{f_y}{n_x}. \tag{A.11}$$

We want to show that there exists a unique equilibrium solution (a_x, b_y) to this system. We are still only interested in the case where $\lambda < 0$ and $\Phi_{xyh} > 0$. Note that we can rewrite the system to obtain expressions that define a_x and b_y as explicit functions of market tightness θ_{xyh}:

$$a_x = \left(\sum_y \sum_h \exp\left((1+\theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda}\Phi_{xyh}\right)\right)^{-1} \quad \text{for all } x,$$

$$b_y = \left(\sum_x \sum_h \exp\left(\theta_{xyh}^{\lambda-1}(1+\theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda}\Phi_{xyh}\right)\right)^{-1} \quad \text{for all } y.$$

Thus, if a solution θ^* exists that satisfies (A.11), we can evaluate these functions to obtain values for a_x and b_y. Since market tightness measures the number of searching firms over workers, we consider only solutions on the strictly positive interval. We thus look for a root of:

$$G(\theta_{xyh}) = \theta_{xyh} - \exp\left(\frac{\theta_{xyh}^{\lambda-1}-1}{(1+\theta_{xyh}^{\lambda})^{(\lambda-1)/\lambda}}\Phi_{xyh}\right)\frac{a_x}{b_y}\frac{f_y}{n_x}. \tag{A.12}$$

The limits of (A.12) on the relevant domain are:

$$\lim_{\theta_{xyh}\to 0} G(\theta_{xyh}) = -\exp(\Phi_{xyh}) \quad \text{and} \quad \lim_{\theta_{xyh}\to\infty} G(\theta_{xyh}) = \infty.$$

The intermediate value theorem assures that at least one root exists, because $G(\theta_{xyh})$ is continuous on the positive strictly interval. We now check whether the function is monotonous to assure uniqueness. The derivative can be expressed (after some rewriting) as:

$$G'(\theta_{xyh}) = 1 - \Phi_{xyh} \frac{\theta_{xyh}^{\lambda-1} \left(\theta_{xyh}^{-\lambda} \left(\theta_{xyh}^{\lambda} + 1\right)\right)^{\frac{1}{\lambda}} (\lambda\theta_{xyh} + \lambda - \theta_{xyh} - 1)}{\theta_{xyh}^{2\lambda} + 2\theta_{xyh}^{\lambda} + 1} z,$$

$$\text{where: } z = \exp\left[\frac{\Phi_{xyh} \left(\theta_{xyh}^{-\lambda} \left(\theta_{xyh}^{\lambda} + 1\right)\right)^{\frac{1}{\lambda}} \left(-\theta_{xyh} + \theta_{xyh}^{\lambda}\right)}{\theta_{xyh}^{\lambda} + 1}\right].$$

The sign of the second term of $G'(\theta_{xyh})$ is always the inverse sign of the surplus Φ_{xyh}, because all components of the second term are positive, except $(\lambda\theta_{xyh} + \lambda - \theta_{xyh} - 1)$, which is always negative under the parameter restrictions. Since the surplus is assumed positive, $G'(\theta_{xyh}) > 0$ and $G(\theta_{xyh})$ is strictly monotone.

Consider now the equations (A.9) and (A.10) as interdependent mappings in (a_x, b_y):

$$F_x(a_x, b_y) = \left(\sum_y \sum_h \exp\left((1 + \theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda} \Phi_{xyh}\right)\right)^{-1} \quad \text{for all } x,$$

$$F_y(a_x, b_y) = \left(\sum_x \sum_h \exp\left(\theta_{xyh}^{\lambda-1} (1 + \theta_{xyh}^{\lambda})^{(1-\lambda)/\lambda} \Phi_{xyh}\right)\right)^{-1} \quad \text{for all } y,$$

$$\text{with: } F_{\theta_{xyh}}(a_x, b_y) : \theta_{xyh} - \exp\left(\frac{\theta_{xyh}^{\lambda-1} - 1}{(1 + \theta_{xyh}^{\lambda})^{(\lambda-1)/\lambda}} \Phi_{xyh}\right) \frac{a_x}{b_y} \frac{f_y}{n_x} = 0.$$

We show that for an arbitrary θ^* the mappings F_x and F_y are contractions. Note that $F_{\theta_{xyh}}(a_x, b_y)$ is monotone decreasing in a_x and increasing in b_y. This implies that $F_x(a_x, b_y)$ is decreasing in a_x and increasing in b_y. The opposite is true for $F_y(a_x, b_y)$, which increases in a_x and decreases in b_y. This shows monotonicity. The system is also contractionary. [incomplete, to finish]

A.3.1 Identification

This Appendix formalizes the identification argument discussed in section **??**. We still consider the setting where we have samples on equilibrium matching patterns and wages in multiple segmented markets z. Note that we can express the observed matching patterns and wages in terms of their nonlinear model expression as:

$$\hat{\mathcal{M}}_{xyhz} = S_{xyhz} m(\theta_{xyhz}) = \exp\left((1+\theta^{\lambda}_{xyhz})^{(1-\lambda)/\lambda}\Phi_{xyh}\right) a_{xz}\hat{n}_{xz}(1+\theta^{\lambda}_{xyhz})^{1/\lambda}, \tag{A.13}$$

$$= D_{xyhz} q(\theta_{xyhz}) = \exp\left(\theta^{\lambda-1}_{xyhz}(1+\theta^{\lambda}_{xyhz})^{(1-\lambda)/\lambda}\Phi_{xyhz}\right) b_{yz}\hat{f}_{yz}(1+\theta^{\lambda}_{xyhz})^{1/\lambda}\theta^{-1}_{xyhz}, \tag{A.14}$$

$$\hat{\mathcal{M}}_{x0z} = S_{xyhz}(1-m(\theta_{xyhz})),$$

$$= \exp\left((1+\theta^{\lambda}_{xyhz})^{(1-\lambda)/\lambda}\Phi_{xyh}\right) a_{xz}\hat{n}_{xz}\left(1-(1+\theta^{\lambda}_{xyhz})^{1/\lambda}\right), \tag{A.15}$$

$$\hat{\mathcal{M}}_{0yz} = D_{xyhz}(1-q(\theta_{xyhz})),$$

$$= \exp\left(\theta^{\lambda-1}_{xyhz}(1+\theta^{\lambda}_{xyhz})^{(1-\lambda)/\lambda}\Phi_{xyh}\right) b_{yz}\hat{f}_{yz}\left(1-(1+\theta^{\lambda}_{xyhz})^{1/\lambda}\theta^{-1}_{xyhz})\right), \tag{A.16}$$

$$\hat{\mathcal{W}}_{xyhz} = \frac{\rho_{xyh} - \theta^{\lambda}_{xyhz}\alpha_{xyh}}{1+\theta^{\lambda}_{xyhz}}. \tag{A.17}$$

To prove identification, we show that these non-linear equations are monotonic (and can thus be inverted) in the parameters α_{xyh}, ρ_{xyh}, and λ.

The Joint Surplus. The joint surplus $\Phi_{xyh} = \alpha_{xyh} + \rho_{xyh}$ can be identified from the number of matches $\hat{\mathcal{M}}_{xyhz}$ increasing monotonically in the surplus. This can be shown by simply rewriting either equation (A.13) or (A.14) as a linear function in $\exp(\Phi_{xyh})$. Note that we do not need to consider how the surplus recursively enters the market tightness function, which can be shown to be monotonic through the same argument.

Amenities and Productivity. Separate identification of the amenity and productivity terms, given the joint surplus, follows immediately from the observation of wages. We can see from equation (A.17) that, for positive market tightness values, equilibrium wages increase in productivity and decrease in amenities.

The Matching Function Elasticity. To show identification of λ we can rely on the proof in

Beauchamp et al. (2024), who show monotonicity of the same matching function in λ. Note that with observations on matching patterns and wages in a single market, the amenity and productivity terms would exhaust all data moments. With multiple markets, we can exploit variation in matching patterns and wages across markets to identify additional parameters. Note that, since we have only a single parameter in the matching function, observations in two markets suffices for identification.

A.4 Data

A.4.1 Job Vacancies

This appendix describes in detail how to construct and scale the distribution of job vacancies over occupations for each state and year. We rely on two sources of data, which respectively inform us about the total number of job vacancies, and the occupational distribution:

1. *Vacancies: Levels.* The Job Openings and Labor Turnover Survey (JOLTS) collected by the Bureau of Labor Statistics (BLS) contains monthly data on the number of job openings and hires at more than 20,000 establishments in the United States.[13] This data is aggregated into state-level time series on the number of job openings and the number of job openings *per unemployed worker*.

2. *Vacancies: Occupational Distribution.* The National Labor Exchange (NLx) Research Hub job vacancy database contains more than 150 million unique job postings. These vacancies identify several job characteristics – of which we mainly use the state of employment and the occupation. We discuss more features of the data in section A.4.2.

To construct the number of job vacancies at the state $\times$ occupation level, we first estimate the *distributions* with the NLx vacancy data. We scale these distributions by the number of job vacancies *per unemployed worker* (in the JOLTS) and the number of unemployed workers in our sample (in the ACS) to obtain a final scaled count of vacancies for our sample.

A.4.2 The National Labor Exchange (NLx) Research Hub Database

We use a database on job vacancies provided by the Research Hub of the National Labor Exchange (NLx). The NLx is a non-profit online labor exchange operated by the National Association of State Workforce Agencies (NASWA), which represents the workforce agencies of all states in the United-States, and the DirectEmployers Association (DE). The vacancy data provided by the NLx Research Hub is collected from several sources: companies' career websites, states' job

[13]A job opening in JOLTS is defined based on three conditions that need to be satisfied *on the last business day of the month*: (i) a position exists and has work available, (ii) the job could start within thirty days, and (iii) there is active recruiting for workers from outside of the establishment.

vacancy banks, and the United States' federal jobs portal. These vacancies are vetted to remove duplicates and junk records. The data is available in real time, and historical snapshots have been archived for almost fifteen years. These snapshots provide information on all vacancies that were open on any given day, which allows us to select only those vacancies open on the last business day of the month (as in JOLTS) to construct the occupational vacancy distributions.

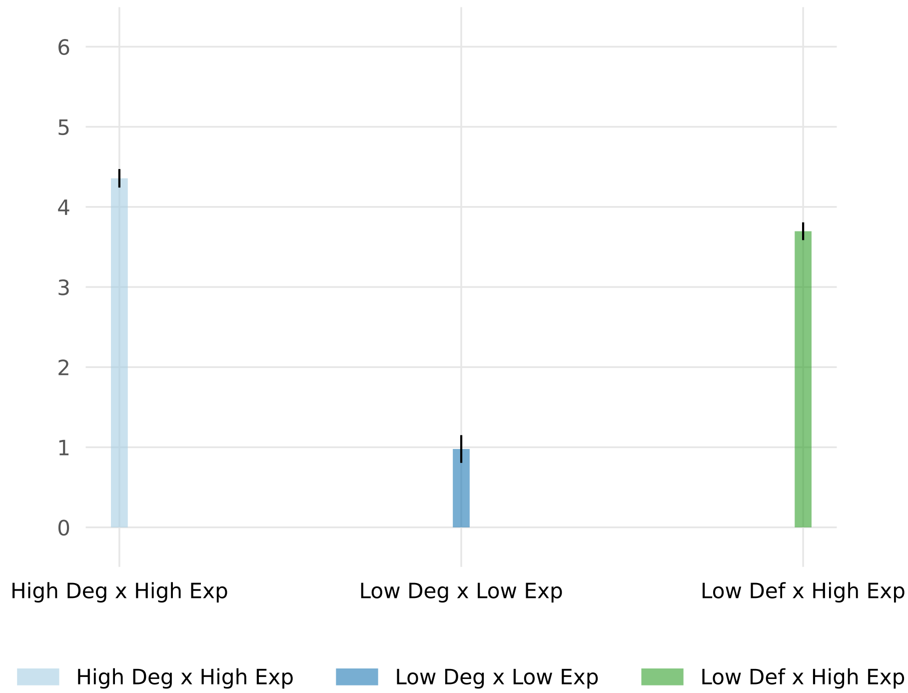
Coverage. The NLx data aims to cover all real job postings in the United States. Table **??** shows that, in recent years, almost XX% of the vacancies in the United States as estimated from JOLTS are represented by an NLx vacancy.[14] This is comparable to the data offered by lightcast.io that has been used in Hershbein and Kahn (2018), Deming and Noray (2020), and Acemoglu et al. (2022). These papers have furthermore argued that online job vacancies are relatively representative in terms of occupations and industries. The main caveat, that high skill jobs are overrepresented and low skill jobs underrepresented, is also less important because of our focus on higher earning occupations.

[14] But note that coverage in earlier years is significantly worse.

A.5 Results

This appendix contains further results from the structural model.

Figure 7: Productivity Across Worker Types



Notes. Productivity estimates of different worker types – averaged across gender.