DISCUSSION PAPER SERIES

DP17469

Labor Supply and Firm Size

Lin Shao, Faisal Sohail and Emircan Yurdagul

MACROECONOMICS AND GROWTH

ISSN 0265-8003

Labor Supply and Firm Size

Lin Shao, Faisal Sohail and Emircan Yurdagul

Discussion Paper DP17469
Published 13 July 2022
Submitted 13 July 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Labor Supply and Firm Size

Abstract

Larger firms exhibit i) longer hours worked, ii) higher wages, and iii) smaller (larger) wage penalties for working long (short) hours. We reconcile these patterns in a general equilibrium model, which features the endogenous interaction of hours, wages, and firm size. In the model, workers willing to work longer hours sort into larger firms that offer a wage premium. Complementarities in hours generate wage penalties that increase with the distance from average firm hours. We use the model to argue the importance of the interaction between hours, wages, and firm size for inequality and firm policy.

JEL Classification: E23, J22, J23, J31

Keywords: Labor Supply, hours worked, firm size, sorting, complementarities

Lin Shao - lin.j.shao@gmail.com
Bank of Canada

Faisal Sohail - faisal.sohail@unimelb.edu.au
University of Melbourne

Emircan Yurdagul - eyurdagu@eco.uc3m.es
Universidad Carlos III de Madrid and CEPR

Acknowledgements
The views expressed in this paper are solely those of the authors and may differ from the official Bank of Canada views. We thank Alexander Bick, Andres Erosa, Belen Jerez, Minjoon Lee, Lawrence Uren, Felix Wellschmied, and seminar participants at CEA 2022, Bank of Canada, Singapore Management U. Macro Workshop, Trinity College Dublin, U. Alicante, U. Catolica Uruguay, U. Helsinki, U. Houston, U. Melbourne, and VAMS for valuable feedback. Yurdagul gratefully acknowledges the grant IJC2018-038229-I by MCIN, grant 2020/00369/001 by AEI, grants EPUC3M11 (V PRICIT) and H2019/HUM-589 by Comunidad de Madrid.

Labor Supply and Firm Size*

Lin Shao† Faisal Sohail‡ Emircan Yurdagul§

Bank of Canada U. Melbourne U. Carlos III and CEPR

June 2022

Abstract

Larger firms exhibit i) longer hours worked, ii) higher wages, and iii) smaller (larger) wage penalties for working long (short) hours. We reconcile these patterns in a general equilibrium model, which features the endogenous interaction of hours, wages, and firm size. In the model, workers willing to work longer hours sort into larger firms that offer a wage premium. Complementarities in hours generate wage penalties that increase with the distance from average firm hours. We use the model to argue the importance of the interaction between hours, wages, and firm size for inequality and firm policy.

JEL Codes: E24, J2, J31

Keywords: labor supply, hours worked, firm size, sorting, complementarities

*The views expressed in this paper are solely those of the authors and may differ from the official Bank of Canada views. We thank Alexander Bick, Andres Erosa, Belen Jerez, Minjoon Lee, Lawrence Uren, Felix Wellschmied, and seminar participants at CEA 2022, Bank of Canada, Singapore Management U. Macro Workshop, Trinity College Dublin, U. Alicante, U. Catolica Uruguay, U. Helsinki, U. Houston, U. Melbourne, and VAMS for valuable feedback. Yurdagul gratefully acknowledges the grant IJC2018-038229-I by MCIN, grant 2020/00369/001 by AEI, grants EPUC3M11 (V PRICIT) and H2019/HUM-589 by Comunidad de Madrid.

†lin.j.shao@gmail.com. Bank of Canada, 234 Wellington Street, Ottawa, ON K1A 0G9, Canada.

‡faisal.sohail@unimelb.edu.au. 111 Barry St, Carlton VIC, 3053, Australia

§emircan.yurdagul@uc3m.es. Calle Madrid 126, Getafe-Spain 28903.

1 Introduction

There exists significant variation in workers' hours across firms. This heterogeneity in labor supply has important implications for inequality, particularly if variation in hours is related to wages and firm characteristics. For example, if workers in high-paying firms also work longer, variation in hours amplifies income inequality, while inequality is mitigated if workers in high-paying firms work fewer hours. Despite its importance, macroeconomic models of production typically abstract from studying the distribution of hours across firms and, as a result, neglect the interaction between hours, wages and firms.

This paper studies the relationship between hours, wages and firms. Empirically, we document a systematic relationship between wages, hours and firm size. Motivated by this relationship we develop a general equilibrium framework with heterogeneous workers and firms in which production exhibits complementarities in working hours and hence depends on the distribution of workers' hours within a firm. The model equilibrium features heterogeneity in working hours across firms and an endogenous interaction of hours, wages, and firm size that is consistent with the patterns we highlight.

Using this theoretical framework, we argue that incorporating the relationship between hours, wages and firm productivity into models of firm production has important aggregate implications for earnings inequality and firm policy – implications that are absent in canonical models that abstract from this relationship. Specifically, we show that in both the data and model, heterogeneity in hours contributes significantly to inequality in wages and incomes even after controlling for worker and firm characteristics, and this contribution depends on the degree of complementarity in working hours. In addition, we highlight a novel channel through which size-dependent firm policies generate welfare losses for workers by shrinking the number of options, in terms of hours and wages, that are available to heterogeneous workers.

The first part of the paper uses data from the US Current Population Survey (CPS) to document three motivating facts on the relationship between hours and wages by firm size. We begin by documenting a relatively understudied empirical pattern; average hours worked increase with firm size. Second, we revisit the well-documented size-wage premium – that is, average hourly wages increase with firm size (Brown and Medoff 1989, Oi and Idson 1999). Finally, our third fact is novel. We find that wage penalties for working relatively long and short hours differ by firm size. In particular, we show that workers in larger firms exhibit smaller wage penalties for working long hours but larger wage penalties for working short hours.[1]

Motivated by this evidence, the second part of the paper discusses the theoretical framework we develop to study the interaction of hours, wages, and firms. In the model, firms differ in their exogenous productivity and decide on static labor input.[2] Consistent with the findings in Shao et al. (2022) who argue that working hours are complements in production, the firm's production function allows for complementarities between workers' hours. Specifically, the labor input of firms is a non-linear aggregate of the hours worked by all workers such that workers are more productive if they work similar hours as their co-workers. Workers vary in their value of leisure, labor efficiency or skills and have preferences for working in firms of different productivities. Given these three dimensions of exogenous heterogeneity, workers choose their labor supply and which firm to work for.

Despite its minimal structure, the model which is calibrated to match key features of the US economy, replicates all three motivating facts. First, the size-wage premium is generated by an interaction of firm-level heterogeneity in productivity and workers' preferences over working in firms of different characteristics. As in Card et al. (2018), with such idiosyncratic tastes, workers view firms of different productivity as imperfect substitutes. This makes firm employment increase less steeply with productivity and allows the marginal productivity of labor (wages) to increase with firm productivity, hence size, in equilibrium. We introduce heterogeneity in tastes as a simple and relatively standard feature to deliver a size-wage premium which interacts with other mechanisms in the model to generate our other two motivating facts. Importantly, this modeling choice for generating the size-wage premium does not, on its own, generate increasing hours with firm size, nor the size-dependent wage

[1]In the Appendix, we also document our findings for firms and establishments using the Labour Force Survey from Canada. Our empirical findings are also evident at both the firm and establishment level in Canada.

[2]Several papers document a robust positive correlation between measures of firm productivity and the number of employees in a firm (see, for example, Leung et al., 2008 and Bartelsman et al., 2013). Consistent with this, we use the number of employees in a firm to proxy for productivity.

penalties for short and long hours. Indeed, as we discuss below, differences in firm productivity and endogenous sorting between workers and firms remains critical for reconciling the data.

Second, a positive relationship between firm size and worker hours is driven by the interaction of size-wage premium and workers' labor supply decisions, including worker sorting into firms of different sizes and their choice of hours worked in a given firm. The size-wage premium introduces two opposing forces that determine worker sorting into firms of different sizes. Long-hour workers, for instance, those with a low value of leisure, enjoy larger income gains by working for large (high wage) firms. On the other hand, short-hour workers have lower consumption and higher marginal utility; therefore, they value the additional income more than longer-hour workers. Conditional on the sorting decision, two similar and opposing forces determine whether a worker works longer hours in large firms than in small firms. Higher wages in large firms incentivize workers to work longer because they generate greater income gains from extra hours worked. On the other hand, the marginal utility of the same worker is higher in small firms due to a lower income, which encourages them to work longer hours. The relative magnitude of these two sets of opposing forces determines the net effects of the two margins of labor supply decisions. In our calibrated model, the first effect dominates in both cases. In equilibrium, workers with longer desired hours select into larger firms, and for any given type, workers also work longer hours in large firms. Therefore, average working hours increase with firm size.

Third, complementarities in workers' hours generate short- and long-hour wage penalties for deviating from the usual hours worked in a firm. Workers' productivity declines as the gap between their hours and the firm's usual working hours widens, and consequently, they suffer more significant wage losses. As larger firms feature longer average hours worked, long (short) hours are less (more) heavily penalized compared to small firms.

The model also has strong implications for workers sorting based on their desired hours. Specifically, it predicts that workers that work fewer hours than their coworkers are more likely to sort into smaller firms where their desired hours are closer to their average coworkers. The converse is true for workers that work longer hours than their coworkers. We find support for this prediction in the data. By tracking CPS respondents over 12 months, we show that workers with relatively shorter hours are more likely to transition into smaller firms and less likely to transition into larger firms – as predicted by the theory. This finding complements existing literature studying labor market sorting by highlighting the importance of sorting based not only on worker skills and firm heterogeneity but also workers' desired hours.

After detailing the model mechanisms that reconcile our motivating facts, we study the aggregate implications of these mechanisms. First, we explore how heterogeneity in working hours shapes wage inequality in the data and model. Consistent with the data, the model features wage dispersion due to heterogeneity in worker skill, firm productivity, *and* working hours. Although the importance of worker and firm heterogeneity for wage inequality is well-studied, the role of hours heterogeneity is not. We find that heterogeneity in hours accounts for 13% of *explained* wage dispersion, compared to a 19% contribution observed in the data. Through the lens of the model, working hours impact wage dispersion due to complementarity in workers' hours. Indeed, without complementarities, wages within firms would be independent of working hours, and hours heterogeneity would not influence wage inequality, conditional on firm productivity.

We next explore the role of hours on inequality in income – the product of hours and wages. A variance decomposition exercise reveals that hours' dispersion and covariance with wages account for around 20% of overall income dispersion in the data and model. In addition, we highlight the role of complementarity in working hours in influencing income inequality. Although complementarities generate wage dispersion, they also introduce incentives for workers to work similar hours through wage penalties for working dissimilar hours, compressing the distribution of hours in the economy. Quantitatively, we find that the latter channel dominates, and income inequality decreases in the degree of complementarity in hours.

Finally, we explore the implications of size-dependent firm policies on worker welfare using our theoretical framework. While it is well-established that these policies are distortionary and lead to aggregate productivity and output losses (see, for example, Guner et al. 2008), we highlight a novel channel through which complementarities in hours might also impact worker welfare in response to such firm policies. Specifically, we compare the impact of size-dependent distortions in our benchmark model to a version of the model without complementarities. Despite featuring similar output losses, we find that the benchmark model generate much greater welfare losses than the model without complementarities. In the economy featuring complementarities in hours, workers with short or long hours are penalized by deviating from firms' usual hours worked. However, workers can mitigate these losses by sorting into small or large firms with average hours more similar to their own. Distortions shrink the dispersion in net productivity of firms and make the average wage — hence average hours worked — more similar across firms. As a result, workers have fewer options to choose from in terms of average firm hours. In our quantitative exercise, the dampened sorting of hours across firms particularly affects high-leisure (short-hour) workers. This analysis shows that common policies such as subsidies to small firms or more onerous regulations for larger firms can adversely impact workers – an insight that policymakers may overlook if they ignore the endogenous interaction of hours, wages, and firm size.

Taken together, this paper highlights the endogenous interaction between hours, wages, and firm-level heterogeneity. An interaction that is currently under-emphasized in the literature but, as we argue, has important implications for aggregate outcomes.

Related literature This paper is closely related to several strands of literature studying the interaction of firm characteristics, wages and hours worked. Among our three motivating facts, the size-wage premium has been most extensively documented and studied. There is no consensus on an exact determinant of the size-wage premium, and we generate a size-wage premium through workers' heterogeneous preferences over potential employers, closely following the approach in Card et al. (2018) and Lamadon et al. (2022). Such heterogeneity prevents worker flows from equalizing wages across firms and generates higher wages for more productive (larger) firms. Moreover, high-efficiency workers are attracted by the higher wages offered by larger firms, so our model also features positive sorting of high-skill workers into high-productivity firms. Such positive sorting is frequently highlighted as contributing to the size-wage premium and is evident in our framework (see, for example, Oi and Idson, 1999).

To our knowledge, there are two empirical studies – Montgomery (1988) and Headd (2000) – that touch on the relationship between hours worked and firm size. Both studies show that larger firms have a lower fraction of workers working part-time. These facts are related to our empirical findings, where we show that the average hours worked increases with firm size. Unlike these papers, we compare the full distribution of hours across establishments of different sizes, moving beyond the distinction between part-time and full-time workers. In line with this, our model is rich enough to capture differences between small and large firms that are not driven by discontinuities in costs of hiring part and full-time workers or workers' productivity.

Our third fact, the variation in the long and short hours penalties across size categories of firms, is novel and naturally relates to the literature that documents the presence of such penalties across the economy. Recent work by Yurdagul (2017) and Bick et al. (2020) have documented a hump-shaped relationship between wages and hours in the aggregate while Shao et al. (2022) document such a relationship within establishments. We build on these findings by providing new evidence showing that the hump-shaped wage-hours profile varies with firm size.

Our theoretical framework joins a growing literature studying heterogeneous agent macroeconomics models that aim to be consistent with micro-level evidence on wages and the labor supply of workers. Much of this literature has focused on the response of labor supply to business cycle or life-cycle fluctuations (see for example, Heathcote et al. 2014, Erosa et al. 2016, and Chang et al. 2020). Instead, we focus on the cross-sectional relationship between hours, wages, and firm characteristics. Bick et al. (2020) study the relationship between hours and wages. They present a structural model of labor supply and earnings that replicates the hump-shaped wage-hours profile via an exogenously specified non-linear wage schedule. In contrast, we endogenously generate a non-linear wage schedule in a general equilibrium model. Further, we focus on the interaction between hours, wages and firm-level heterogeneity, and show that this interaction has important implications for inequality and welfare.

A distinguishing feature of our model is the presence of complementarities in the hours of different workers. We introduce this feature to capture the idea that workers must coordinate their work schedules to produce output. Importantly, complementarity in hours is consistent with empirical evidence. Recent work by Shao et al. (2022) uses matched employer-employee data to document that workers' hours within the same establishment are gross complements. Our modeling of complementarities is most closely related to Yurdagul (2017) which studies the flexibility motive behind entrepreneurship.[3] Complementarity in hours can limit workers' ability to choose their working hours. Hence, our model relates to the literature exploring the impact of constraints on working hours resulting from coordination (see, for example, Altonji and Paxson 1988, Chetty et al. 2011, Labanca and Pozzoli 2021, and Cubas et al. 2019). The importance of flexibility in choosing one's hours has also been emphasized in the literature that aims to understand the gender differences in occupational choices (Erosa et al., 2016 and Bento et al., 2021). We contribute to these strands of literature by explicitly incorporating inflexible working hours (via complementarities) in a theoretical model and studying its interaction with firm heterogeneity and its implications for shaping the sorting pattern of workers across firms.

Our analysis of wage inequality builds on empirical work such as Song et al. (2019) and Barth et al. (2016) that explore the role of worker and firm characteristics in generating wage dispersion. We contribute to this literature by documenting the role of hours heterogeneity for wage inequality. Blau and Kahn (2011) and Checchi et al. (2016) document the relationship between heterogeneity in hours and income inequality. We relate to this literature by highlighting the importance of complementary in hours. There is extensive study on the impact of size-dependent firm policies. Much of this literature focuses on the impact of aggregate outcomes such as output, productivity and the firm size distribution (see, for example, Guner et al. 2018, Gourio and Roys 2014, and Bento and Restuccia 2017). Our quantitative analysis of size-dependent policies follows this literature and documents a novel channel by which worker welfare may be impacted since these policies shrink the options available to workers in terms of hours worked.

An outline of the paper is as follows. Section 2 describes our motivating facts in detail and Section 3 outlines our model. In Section 4 we calibrate the model and evaluate its quantitative predictions. Section 5 explores the aggregate implications of our theoretical framework for wage and income inequality, and size-dependent firm policies. Section 6

[3] Battisti et al. (2021) use a similar production function to discuss the link between estimates of the Frisch elasticity and degree of complementarities in hours.

concludes.

2 Motivating Facts

This section documents three motivating facts about the distribution of hours and wages across firm size. First, we establish a robust positive relationship between firm size and average worker hours. Workers in the smallest firms (under 10 employees) work 7% fewer hours per week than those in the larger firms. Second, we confirm the existence of a size-wage premium. Third, we show that workers face penalties for working either short or long hours in all firms. However, the magnitude of these penalties vary systematically across size categories. In particular, we find that the penalty for working long hours decreases with size while the penalty for working short hours increases with size.

Data description To establish these facts, we use data from the Annual Social and Economic Supplement (ASEC) of the CPS covering information from 1991 to 2018. This supplement to the CPS contains detailed information on respondents' economic activity for the past year. Importantly, it includes information on worker earnings, usual weekly hours worked, and firm size.[4] The partitioning of size bins has varied over time, so for clarity and consistency, we report three categories of firm size; i) small (under 10 employees), ii) medium (between 10 and 100 employees), and iii) large firms (over 100 employees). We restrict attention to individuals aged between 25 and 64, who worked with a single private employer in the previous year and exclude those who usually work less than 10 hours a week or earn less than half the federal minimum wage. Respondents with imputed values for firm size, hours worked, or weeks worked are also dropped. The final sample includes just over 1 million respondents.[5]

[4]Data is extracted from IPUMS and described in Flood et al. (2020). Firm size is reported in bins and records the total number of employees that a worker's employer has at all establishments. For example, the 2019 ASEC, which contains information for the reference year 2018, asks the following: "Counting all locations where this employer operates, what is the total number of persons who work for your employer?". We start our sample in 1991 since, prior to this, the smallest reported firm size category was firms with under 25 employees.

[5]Analogous data for Canada from the 1997 to 2018 Labour Force Surveys (LFS) is presented in the Appendix B.2 for both establishments and firms.

Fact 1 Average hours increase with size.

We begin by studying the relationship between firm size and hours worked. Figure 1 reports the distribution of usual weekly hours worked by firm size with Panel (a) showing the overall distribution of usual hours worked. While the median weekly hours across all firms is between 40 and 44 hours, there are important differences in the share of short and long hours worked across firm sizes. This is evident in Panels (b) and (c), which report, respectively, the distribution of the right and left tails of the hours distribution. Panel (b) shows that workers in small firms are much more likely to work shorter (<40) hours than their counterparts in larger firms. For example, around 3% of employees in larger firms work 30 to 35 hours while the analogous share in small firms is around 6%. Panel (c) shows that employees in medium and large firms are more likely to work between 45-59 hours with a similar likelihood of working very long (≥ 60) hours. For example, just around 8% of employees in small firms (less than 10 employees) work between 50 and 55 hours while the analogous share in larger firms is 10%.

(a) Overall (b) Short hours (c) Long hours

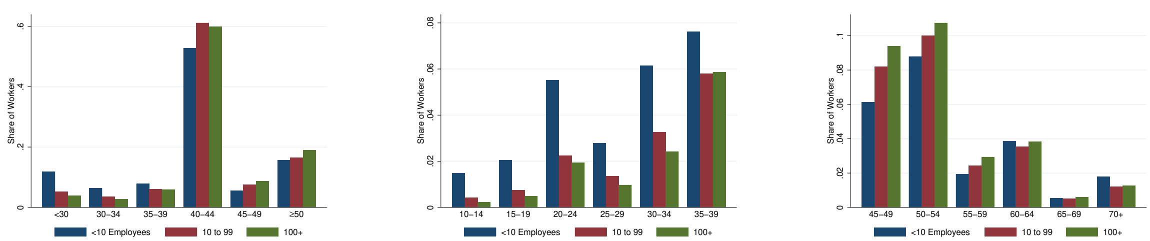

Figure 1: Distribution of working hours by firm size

Notes: The figure reports the share of workers by their usual weekly hours worked and firm size using data from the CPS.

As suggested from Figures 1, and confirmed in Figure 2, there is a positive relationship between average hours worked and firm size. On average, workers in the largest firms work around 3 hours longer than workers in the smallest firms.

While informative, these cross-sectional averages do not control for confounding factors, such as industry, that might impact both firm size and hours worked. To control for such factors, we estimate the following regression,

$$\log(h_i) = \alpha + \left(\sum_{f \in F} \beta_f \mathbb{I}_{i,f}\right) + \delta X_i + \epsilon_i \qquad (1)$$

where h_i is the usual weekly hours worked. X_i is a vector of individual-level controls which includes a quadratic in years of experience, dummies for the race, education, gender, marital status as well as state, year, and industry fixed effects. The variable $\mathbb{I}_{i,f}$ is an indicator variable which is equal to one if an individual is employed in a firm of size $f \in F$ so that the coefficient β_f captures the elasticity of hours worked by firm size.

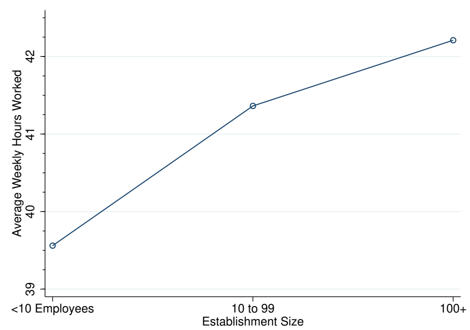

Figure 2: Average weekly hours by firm size

Notes: The figure reports the average usual weekly hours worked by firm size using data from the CPS.

Table 1 reports this elasticity and shows that when excluding controls for industry or demographics, workers in larger US firms work around 6% to 9% longer hours than workers in firms with under 10 employees. Controlling for industry and demographic characteristics explains some of the differences in hours worked between medium and larger firms and implies that workers in firms with over 10 employees work between 5% and 7% longer than workers employed in firms with less than 10 employees.

Fact 2 Average wages increase with size.

The wage premium in large firms and establishments has been studied extensively (see, for example, Oi and Idson 1999). We establish the size-wage premium in our data by estimating the following regression,

Table 1: Firm size and hours worked

	(1)	(2)	(3)
10 to 99 Employees	0.063***	0.056***	0.052***
	(0.001)	(0.001)	(0.001)
100+ Employees	0.086***	0.077***	0.069***
	(0.001)	(0.001)	(0.001)
Year, State FE	Y	Y	Y
Demographic Controls	N	Y	Y
2-digit Industry FE	N	N	Y
N	1,000,820	1,000,820	1,000,820
R^2	0.018	0.101	0.113

Notes: The table reports the coefficient β_f estimated from Equation (1) where the reference size category is firms with under 10 employees. Data is from the CPS sample. Standard errors are reported in parentheses. ***, **, * indicates statistical significance at the 1% level.

$$\log(w_i) = \alpha + \left(\sum_{f \in F} \beta_e \mathbb{I}_{i,f} \right) + \left(\sum_{h \in H} \gamma_h \mathbb{I}_{i,h} \right) + \left(\sum_{f \in F} \sum_{h \in H} \theta_{f,h} \left[\mathbb{I}_{i,f} \times \mathbb{I}_{i,h} \right] \right) + \delta X_i + \epsilon_i \tag{2}$$

where $\log(w_i)$ is the log hourly wages of individual i. Hourly wages are computed as the ratio of annual earnings, usual weekly hours and weeks worked. As in (1), X_i is a vector of individual-level controls which includes demographic controls, state, year, and industry fixed effects. The indicator variable $\mathbb{I}_{i,f}$ is equal to one if an individual is employed in a firm of size $f \in F$. Similarly, $\mathbb{I}_{i,h}$ is equal to one if an individual works h hours.

We partition weekly hours into a set H by grouping hours in 5-hour bins. The partitioned set is $H = \{10 - 14, 15 - 19, \ldots, 65 - 69, 70 - 99\}$. The final bin $70 - 99$ is larger as there are relatively few workers working over 70 hours. As most workers work 40 hours, we choose the category $40 - 44$ as the reference category for hours and omit the coefficients γ_{40} and $\theta_{40,f}$ for all f. The reference size category is firms with under 10 employees.

The regression in (2) extends the specification in Bick et al. (2020) by also controlling for firm size and an interaction term between firm size and usual weekly hour bins. Including these regressors allows us to study i) the size-wage premium (fact 2) and ii) the relationship between hours and wages by firm size (fact 3 discussed below).

The coefficient that captures the firm size wage premium (for workers that work in the 40 hours bin) is β_f. Table 2 reports β_f and shows that it increases monotonically in size. Indeed, the wage premium between the largest and smallest size categories is around 23%.

Table 2: The size-wage premium

	(1)	(2)	(3)
10 to 99 Employees	0.122***	0.106***	0.086***
	(0.003)	(0.003)	(0.003)
100+ Employees	0.317***	0.254***	0.228***
	(0.003)	(0.003)	(0.003)
Year, State FE	Y	Y	Y
Demographic Controls	N	Y	Y
2-digit Industry FE	N	N	Y
N	1,000,820	1,000,820	1,000,820
R^2	0.125	0.348	0.374

Notes: The table reports the coefficient β_f estimated from Equation (2) where the reference size category is the smallest size firms/establishments. That is, firms with under 10 employees. The reference hours bin is $40-44$ hours. Data is from the CPS sample. Standard errors are reported in parentheses. *** indicates statistical significance at the 1% level.

Fact 3 Long-hour (short-hour) penalty decreases (increases) with size.

Next, we study the relationship between wages by hours worked and firm size. Panel (a) of Figure 3 plot the unconditional average hourly earnings of workers by hours worked and firm size. Two salient features are evident from these cross-sectional averages. First, the size-wage premium exists across the entire distribution of working hours. Second, we find a hump-shaped relationship between hours and earnings across all firm sizes. That is, there appears to be a wage penalty resulting from working either longer or shorter hours than the modal hours worked.

Panel (b) reports the same relationship while also controlling for observable characteristics. In particular, it reports the sum of the coefficients γ_h and $\theta_{f,h}$ estimated from (2). This

(a) Unconditional (b) Conditional

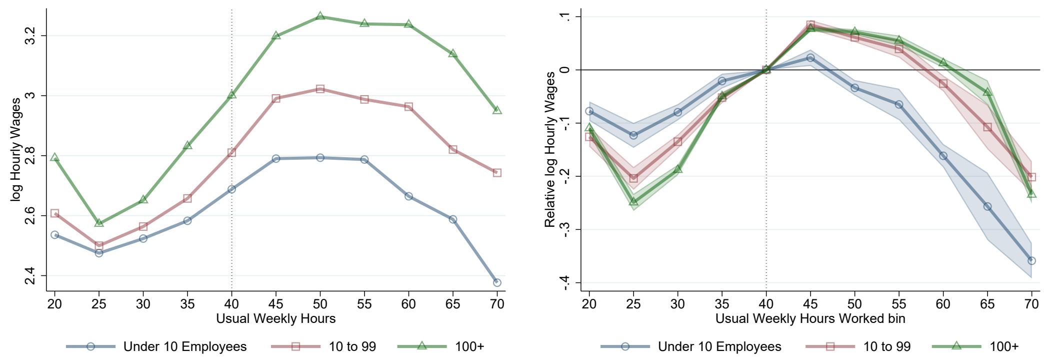

Figure 3: The relationship between wages and hours by firm size

Notes: Panels (a) plots the unconditional average of log hourly wages by usual weekly hours and firm size. Panel (b) reports the sum of coefficients $(\gamma_h + \theta_{f,h})$ estimated from Equation 2. The reference category for hours worked is $40-45$ and the reference category for firm size is firms with under 10 employees. The shaded regions are the 95% confidence intervals. Data is from the CPS sample.

sum captures the wage penalty of working outside of the 40-44 hours bin by size category.[6] An aggregate hump-shaped wage-hours relationship has been documented by Yurdagul (2017) and Bick et al. (2020) in the US. Shao et al. (2022) use a Canadian employer-employee linked dataset to show that wages *within* establishments exhibit a similar hump-shaped relationship with hours. Panel (b) shows that this hump-shaped relationship varies with firm size.

Specifically, the conditional hourly wages in Figure 3 suggest that the penalty for working shorter hours is more severe in larger firms and establishments. In comparison, the penalty for working longer hours is less severe. For example, working around 25 hours in larger firms results in around a 10% *increase* in the wage penalty relative to working the same hours in small firms. On the other hand, working longer hours – say around 60 hours – results in a roughly 15% *decrease* in the penalty relative to working the same hours in small firms.[7]

Together, the three empirical facts highlighted in this section motivate our theoretical analysis. A focus of our framework is the causal link between differential wage penalties

[6]Strictly speaking, the coefficient γ_h captures the wage penalty of working away from the 40 hours bin for small firms. While $\theta_{f,h}$ captures the difference in penalty relative to small firms. Recall that $\theta_{f,h}$ for small firms is zero as small firms are the reference size category.

[7]This difference in relative penalties is captured explicitly by the coefficient $\theta_{f,h}$ in Equation (2). Figure A.1 plots this coefficient and shows that there are indeed statistically significant differences in wage penalties by size. Further, in Appendix B.1, we use the Outgoing Rotation Group of the CPS to control for whether the hourly pay status of workers – that is, whether they earn an hourly wage or are salaried. We find that the composition of hourly earners does not generate qualitatively different implications for wage penalties by firm size.

and average hours across firms. We generate a link between wage penalties and average hours by allowing for complementarities in hours worked in our theoretical model. A natural consequence of such complementarities is that wage penalties are increasing in the distance from the usual hours worked within a workplace. Since larger firms feature longer average hours, longer hours are penalized less severely compared to smaller firms. Conversely, smaller firms feature shorter hours and hence shorter hours are penalized less severely compared to larger firms.[8]

Having discussed our primary motivating facts, we now move to describe our theoretical framework.

3 Model

There are two types of agents in the model economy; firms and households, and we discuss their decisions, in turn, below.

3.1 Firms

There is a continuum of firms with unit mass. For simplicity, we assume that firms' only endogenous input is labor. Production of all the firms in the economy can be represented by $Y = zL^{\theta}$, where L denotes the effective labor input. Firms differ in their exogenous productivity, z, a discrete random variable distribution represented by $F(z)$ which can take J different values. We think of the productivity term z as broadly capturing all non-labor inputs of the firm as well its technology. In what follows, we will denote the index of a firm's productivity level by j and its productivity by z_j.

The effective labor input of a firm depends on the distribution of hours worked by workers in that firm: L denotes the aggregate hours of work in the firm. As in Yurdagul (2017), we

[8] Notice, this intuition implies that average hours not only affect the wage-hours menu faced by workers but are also affected by these wage-hours menus as they alter workers' labor-supply decisions. This feedback mechanism will also be present in our model.

allow for complementarities between hours of workers:

$$L = \left(\int_{i \in N} x_i l_i^\rho di \right)^{\frac{1}{\rho}} \left(\int_{i \in N} x_i di \right)^{1-\frac{1}{\rho}}, \tag{3}$$

where N is the set of workers, and $\{l_i\}_{i \in N}$ is their hours worked. The contribution of each worker to the labor input is scaled up by their efficiency units $x_i > 0$. We abstract from the indices of workers by rewriting the aggregation in terms of the measure of workers employed at each level of hours worked:

$$L = \left(\int_{x \in B_x} \int_0^1 x \mu(l,x) l^\rho dl dx \right)^{\frac{1}{\rho}} \left(\int_{x \in B_x} \int_0^1 x \mu(l,x) dl dx \right)^{1-\frac{1}{\rho}}, \tag{4}$$

where $\mu(l,x)$ is the measure of workers with efficiency x working l hours.

Labor markets are perfectly competitive within each productivity level z_j of firms, with the firms taking the wage schedule $w_j(l,x)$ as given. Given the wage schedule, firms must decide on the measure, $\mu_j(l,x)$, of a given hour-efficiency combination worker to hire in order to maximize their static profits. It will become apparent later that in equilibrium, not all $l-$types might be available on the $j-$market, and we will only derive equilibrium wages for those that are in positive supply. In addition to wages, firms also take as given the available hours in the market z_j, denoted by $[\underline{l}_j, \bar{l}_j]$ in their optimization problems.

Firms in market j maximize their profit by choosing a labor demand schedule μ_j:

$$\pi_j = \max_{\mu_j} Y - \int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} w_j(l,x) \mu_j(l,x) l dl dx \tag{5}$$

$$s.t.\, Y = z_j \left[\left(\int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} x \mu_j(l,x) l^\rho dl dx \right)^{\frac{1}{\rho}} \left(\int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} x \mu_j(l,x) dl dx \right)^{1-\frac{1}{\rho}} \right]^\theta$$

$$\mu_j(l,x) \in [0, \bar{\mu}_j] \;\; \forall l \in [\underline{l}_j, \bar{l}_j].$$

Here we restrict the demand schedule to be between zero and $\bar{\mu}_j$, a scalar that is sufficiently large and positive, for each hour-efficiency pair of workers. The upper bound will be far from binding in the equilibrium but is imposed in order to avoid corner solutions in which each firm hires workers with only one level of hours.[9] The optimal measure of labor is denoted by $\mu_j^*(l, x)$. We assume that firms are held by absentee entrepreneurs that are outside of our model.

3.2 Workers

There is a continuum of infinitely living workers with unit mass. Preferences are given by:

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{c_{it}^{1-\gamma}}{1-\gamma} - v_{it} \frac{l_{it}^{1+\phi}}{1+\phi} \right]. \tag{6}$$

For all workers, the value of leisure follows a Markov process $\Gamma_v(v'|v)$ and idiosyncratic efficiency follows $\Gamma_x(x'|x)$. We denote the set of value of leisure shocks by B_v and the set of efficiency shocks as B_x.

There are two types of shocks directly related to the occupational choice of workers. First, workers have a probability s each period of being able to choose which firm productivity level z_j to work at. With probability $1 - s$, a worker remains in her previous firm productivity group. This rigidity allows the model to generate the persistence of workers in the same size group of firms as in the data.

Second, each period, a worker receives shocks to the value obtained in each firm productivity group. Formally, there is a vector $\boldsymbol{\epsilon} = \{\epsilon_1, \epsilon_2, ..., \epsilon_J\}$ with as many components as the number of different firm-level productivity. This vector follows a distribution $G(\boldsymbol{\epsilon})$ and is drawn independently every period. These shocks capture workers' taste for the other factors locating a worker into a firm not featured in our model and their role is discussed in more detail below.

Once all of the shocks are realized, and the worker knows the firm-level productivity that

[9]This $\bar{\mu}_j-$constraint limits firms' profit under the corner solutions. It helps rule out the corner solutions and allows us to focus on the symmetric equilibrium where firms in the same market demand a uniform share of the total supply of a given worker type. Not only is the symmetric solution tractable, but it is also consistent with the existence of within-firm hours dispersion as observed in the data. As such, we view the $\bar{\mu}_j-$constraint as capturing some aspects missing from our model that prevent firms from hiring only certain hour types, including labor market frictions, regulatory restrictions, heterogeneity in the hours required for each task, and diminishing returns to each hours group, etc.

she will work for, she chooses the labor supply. Accordingly, her value function is:

$$V(a,x,v,\boldsymbol{\epsilon}) = \max_j \{V^G(a,x,z_j,v) + \epsilon_j\}$$

where the value conditional on working in a firm of productivity j is:

$$\begin{aligned} V^G(a,x,z_j,v) &= \max_{a'\geq 0, l\geq 0} \frac{c^{1-\gamma}}{1-\gamma} - v\frac{l^{1+\phi}}{1+\phi} \\ &\quad +\beta E_{x',v',\boldsymbol{\epsilon}'|x,v}\left[sV(a',x',v',\boldsymbol{\epsilon}') + (1-s)\left(V^G(a',x',z_j,v') + \epsilon'_j\right)\right] \\ \text{s.t} &\quad c = w_j(l,x)l + a(1+r) - a'. \end{aligned}$$

Without loss of generality, one can think of the ϵ-shocks as being realized after the value of leisure and the efficiency shock of a worker. The probability of a worker choosing a firm productivity level z_j (conditional on that she has the opportunity to choose) is denoted by $\mathbf{o}(a,x,z_j,v) \in \{1,...,J\}$. Meanwhile, $\mathbf{l}(a,x,z_j,v)$ denotes the labor supply policy function.

Stationary general equilibrium. A stationary general equilibrium consists of a set of policy functions: $\mu_j^*(l,x)$ for firms $j \in \{1,..,J\}$ and $\mathbf{l}(a,x,z_j,v)$ and $\mathbf{o}(a,x,z_j,v)$ for workers in firm group j, wage functions $w_j(l,x)$, and a time-invariant distribution of workers $\varphi(a,x,z_j,v)$ over the type of their employer (z_j), wealth (a), idiosyncratic efficiency x, and the value of leisure v, such that:

(i) The policy functions solve the problems of workers and firms.

(ii) Labor markets clear. The total measure of workers demanded by all firms for each level of firm productivity z_j, idiosyncratic efficiency, x, and working hours $l \in [0,1]$ is equal to the corresponding labor supply:

$$\mu_j^*(l,x)F(z_j) = \int_{a=0}\int_{v\in B_v} \varphi(a,x,z_j,v)\mathbb{1}[\mathbf{l}(a,x,z_j,v) = l]dvda, \ \forall x \in B_x, \ j \in \{1,2,..,J\}$$

(iii) The evolution of the distribution across workers satisfies, for each $a \geq 0$ $x \in B_x$,

$v \in B_v$, and $j \in \{1, 2, .., J\}$:

$$\varphi(a, x, z_j, v) = \int\int\int \Gamma_x(\tilde{x}, x)\Gamma_v(\tilde{v}, v)\mathbf{o}(a, x, z_j, v)\times$$
$$\sum_{\tilde{j}=1}^{J}\varphi(\tilde{a}, \tilde{x}, z_{\tilde{j}}, \tilde{v})\left(s + (1-s)\mathbb{1}[\tilde{j} = j]\right)\mathbb{1}[\mathbf{l}(\tilde{a}, \tilde{x}, z_{\tilde{j}}, \tilde{v}) = a]d\tilde{x}d\tilde{v}d\tilde{a}.$$

Wage schedule. The model equilibrium can only be solved numerically, but we can analytically characterize the wage schedule in equilibrium. We focus on the symmetric equilibrium where firms in each z_j market demand a uniform share of the total supply of a given hour-efficiency type. That is, the demand schedule for every z_j-firm is $\mu_{a,j}(l) \equiv \mu^s_j(l, x)/F(z_j)$ in this equilibrium. In Appendix C.3, we solve for the optimization problem of the firms and show that under fairly relaxed conditions, this demand schedule is optimal for firms. In that case, we can substitute the distribution of (l, x) within the z_j-market to the optimality conditions of labor demand and derive the equation for the equilibrium wages:

$$w_j(l, x) \quad = \quad \theta z_j x L^{\theta-1}_{a,j} E_{a,j}(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^{\rho-1}}{E_{a,j}(l^\rho)} + (1 - \frac{1}{\rho})l^{-1}\right], \tag{7}$$

where on the right-hand side we have the marginal productivity of a worker with l hours and efficiency units of x. In particular, $E_{a,j}(l^\rho) \equiv \big(\int_{x\in B_x}\int_{\underline{l}_j}^{\bar{l}_j} x\mu_{a,j}(l, x)l^\rho dldx\big)/\big(\int_{x\in B_x}\int_{\underline{l}_j}^{\bar{l}_j} x\mu_{a,j}(l, x)dldx\big)$ is the "weighted average" of l^ρ, and $L_{a,j}$ is the labor aggregation implied by the symmetric demand schedule $\mu_{a,j}(l)$. Alternatively, the earnings are given by:

$$w_j(l, x)l \quad = \quad \theta z_j x L^{\theta-1}_{a,j} E_{a,j}(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^{\rho}}{E_{a,j}(l^\rho)} + 1 - \frac{1}{\rho}\right], \tag{8}$$

Equation (7) shows that a worker's wage depends on her hours of work relative to fellow workers in the same firm. The maximum hourly wage is achieved at: $l^*_j = E_{a,j}(l^\rho)^{\frac{1}{\rho}}$. Wages decrease as working hours get further away from l^*_j. As we will discuss in what follows, this will be key in generating not only a hump-shaped wage-hours relationship overall, but also in making the penalties for short and long hours depend on the usual hours in each firm, hence change across firms as we highlight in our motivating facts.

3.3 Discussion

Before discussing our quantitative analysis, we provide more detail on some of the non-standard features of our model and the role they play.

Complementarities in hours. Unlike most macroeconomics models of production, we allow for complementarities in the hours of different workers by aggregating hours in a non-linear manner. Intuitively, such complementarity captures the need for workers to coordinate which can arise naturally when workers' tasks require coordination. The need for such coordination has long been recognized since at least Adam Smith's early discussions of assembly line production. Recently, Bick et al. (2020) and Labanca and Pozzoli (2021) document suggestive evidence supporting the presence of complementarities in hours. More direct evidence is presented in Shao et al. (2022) who use matched employer-employee data from Canada to estimate the elasticity of substitution between hours worked and find that working hours are indeed gross complements in production.

Given the evidence supporting complementarities in working hours, we explicitly incorporate them into our production function. As we discuss below, these complementarities play a crucial role in generating the observed hump-shaped relationship in wages and hours. However, the presence of complementarities do not, on their own, generate differences in wage penalties by firm size or the patterns of sorting and income inequality that we highlight. Indeed, we will specifically highlight how alternative levels of complementarities impact the model's implications in our quantitative analysis below.

Our specific modeling choice assumes firm-wide complementarities between workers, and not within-skill group. The skills of each worker only come into play in the weight they hold in the non-linear aggregation of hours (i.e. bringing the usual hours in the firm closer to higher efficiency workers) and in their contribution to total labor efficiency. An alternative would be to allow for the non-linear aggregation of hours separately within efficiency group. While our results are robust to this alternative, we maintain our baseline approach for two reasons. Economically, our stance is that the coordination between workers captured by this feature might happen between workers from different tasks, occupations, skills.[10] Technically, our modeling choice implies that we only need to solve for a common usual hours for each labor market, whereas the model solution in the alternative approach would require keeping track of usual hours that are specific to each skill group in each labor market.

Structure of the labor market. We assume that markets are segmented by firm type z and within each segment, there is perfect competition for workers between z-type firms. Perfect competition within each market guarantees a direct mapping between wages and the marginal productivity of workers. The assumption of segmentation generates heterogeneous equilibrium wage schedules between firms of different types – as is observed in the data. Without segmentation, perfect competition in an aggregate market would result in a uniform, economy-wide wage schedule. A more general segmentation whereby firms and workers are randomly allocated into an arbitrary number of sub-markets, with perfect competition within each sub-market, would not change the results of our analysis.

Taste shocks. Workers are assumed to have heterogeneous preferences over their workplaces which are captured by the vector ϵ. We introduce these preferences in order to generate a higher wages in larger firms – that is, the size-wage premium. In the absence of a consensus explanation for the size-wage premium, we interpret this heterogeneity in tastes as capturing a number of (non-wage) factors that affect individuals' sorting into firms of different size and productivity. Such factors include differences in non-pecuniary benefits like workplace safety, childcare, or sick leave provision, as well as differences in technology.[11]

Intuitively, the presence of preferences for workplace prevents wages from equalizing across firms of different productivity and generates higher wages in more productive firms that desire to be larger. Our use of taste shocks to generate the size-wage premium follows the approach of recent work such as Card et al. (2018) and Lamadon et al. (2022). Taste shocks have the benefit of being a simple way to generate a size-wage premium and, importantly,

[10] It is for this reason that our primary empirical estimations omit controls for occupations. Having said this, we show in Appendix B.1 that our results are robust to controlling for worker occupations.

[11] We discuss the literature documenting such differences across firms in Appendix C.1.

these shocks do not by themselves generate the other empirical facts that we focus on.[12] If anything, the presence of taste shocks adds noise to the sorting decisions of workers and weakens the positive relationship between firm size and hours. For instance, in the extreme case of very large taste shocks, workers would sort based primarily on their idiosyncratic preferences for workplace resulting in similar hours workers across firm size. As will be clear below, increased similarity in hours worked across firm size will also result in similarity of short and long hour penalties across firm size.

Taste shocks also provide a computational advantage as they effectively 'convexify' the occupational choice decision of workers by introducing randomness. This transforms workers' policy function to a probability between 0 and 1 rather than a binary of 0 or 1. We provide further detail in Appendix C.

4 Quantitative Analysis

This section describes the quantitative implications of the model. We begin by detailing our calibration strategy and then show that the calibrated model can match our three main motivating facts. Following this, we highlight the mechanisms in the model that result in outcomes that are consistent with the data.

4.1 Calibration

We describe the calibration exercise in three parts: the functional forms, parameters calibrated outside the model, and the parameters calibrated targeting features in the data.

Functional forms. Firm productivity is assumed to follow a Pareto distribution with shape parameter λ, and lowest productivity is normalized to 1. This distribution is approximated using 12 grid points. Our results are robust to increasing the number of grid

[12] An alternative approach to generating the size-wage premium would be to allow for complementarities between worker and firm productivity. Such an approach would generate positive sorting between workers and firms so that high productivity workers sort into into high productivity (large) firms. This type of sorting is a widely cited potential reason behind the size-wage premium (see, for example, Oi and Idson (1999)). Introducing such complementarity would sacrifice the parsimony of our model as it would require solving for the optimal matching/sorting function between workers and firms.

points.

We assume that the ϵ-shocks affecting workers' value in each firm follow a Generalized Extreme Value distribution:[13]

$$G(\boldsymbol{\epsilon}) = \exp\left[-\sum_{j=1}^{J} \exp\left(-\frac{\epsilon_j}{\sigma_\epsilon}\right)\right].$$

The parameter σ_ϵ determines the variance of these shocks.

Even though we assume a discrete Markov process for the value of leisure and idiosyncratic efficiency shocks, we parameterize the grids and the evolution of these shocks to resemble AR(1) processes:

$$\log(v_{i,t+1}) = (1-\rho_v)\log(v_0) + \rho_v \log(v_{i,t}) + \xi_{i,t}, \;\; \xi_{i,t} \sim N(0, \sigma_v)$$

$$\log(x_{i,t+1}) = \rho_x \log(x_{i,t}) + \psi_{i,t}, \;\; \psi_{i,t} \sim N(0, \sigma_x)$$

This way, we boil down the corresponding parameters to ρ_v, σ_v and v_0 for the value of leisure, and ρ_x and σ_x for worker efficiency. We follow Tauchen (1986) to map these AR(1) processes to the discrete Markov processes assumed in the model.

Parameters calibrated outside the model. The model is calibrated to match key features of the US data. We assume that a model period corresponds to one year and set the values of several parameters outside the model by using standard values in the literature. As shown in Panel A of Table 4, we set the labor share of output equal to the standard value of 0.67. Following, Erosa et al. (2016), we set the inverse of the Frisch-elasticity of labor supply (ϕ) at 2, and the annual interest rate (r) at 4 percent. We set the constant relative risk aversion (γ) to 2.

The parameter governing the complementarities between working hours, ρ, is set by using results from Shao et al. (2022). Using matched employer-employee data from Canada, Shao et al. (2022) first provide evidence consistent with the presence of complementarities in hours in production and then, using a generalized version of the production in this paper, they

[13] In Appendix C.1, we provide the implication of the ϵ-shocks and the particular distribution we assume for the policy functions of workers.

estimate the substitution parameter ρ to be around -0.5. Based on these results, and with the underlying assumptions that the production technologies adopted by US and Canadian firms are similar, we set the value of ρ to be -0.5.[14]

Parameters targeting features in the data. The remaining parameters are calibrated to match specific targets in the data. To compare model implications to the data, it is useful to construct model-implied firm size categories to match those in the CPS sample – the source of our motivating facts. These size categories are firms with under 10 employees (small firms), firms with 10 to 99 employees (medium firms), and firms with more than 100 employees (large firms). According to the Business Dynamics Statistics (BDS), the fraction of firms in these three categories in 2015 is 77, 21, and 2 percent, respectively. Firms in our model are categorized as "small", "medium" and "large" so that they replicate this distribution – once firms are sorted by size. We then calibrate the shape parameter of the firm TFP distribution, λ, to match the employment share of the largest firm size category. Table 3 summarizes the statistics for the three size categories in the data and the model. In the last column, we provide information on how the average size in the model increases across categories. The fit of our firm grouping to the data is assuring: not only is our firm grouping similar to the data in terms of the percentage of firms in each size category but also in terms of how average employment increases over the size categories.

As will be evident in post-calibration exercises, the level of consumption, hence the level of net-worth, is important in determining the sorting of long- and short-hours workers in the firms of high- and low-wage. Accordingly, we target the median net-worth to labor income ratio with the inter-temporal discount parameter β. Kaplan and Violante (2010) and Erosa et al. (2016), who also adopt the same target for the calibration of the discount factor, estimate this ratio to be 2.5 for the US, which we follow here.[15]

[14] We justify the assumption of similarity between the US and Canadian economies by replicating, in Appendix B.2, our motivating facts using data from the Canadian Labour Force Survey (LFS). We also experimented with an alternative calibration which used only moments from the Canadian economy as targets and found that the mechanisms we highlight do not qualitatively differ.

[15] This estimate is based on years 1989-1992 from the Survey of Consumer Finances, and excludes the top 5 percent of the wealth distribution among workers to avoid the issues due to the undersampling of the richest households. We impose the same restriction when computing the model counterpart of the wealth-income ratio our simulations.

Table 3: Firm size categorization

Data

	Firm Share	Employment Share	Avg. employment (log)
Small	0.77	0.11	0
Medium	0.21	0.23	2.1
Large	0.02	0.66	5.4

Model

	Firm Share	Employment Share	Avg. employment (log)
Small	0.77	0.11	0
Medium	0.21	0.24	2.1
Large	0.02	0.65	5.4

Notes: The size categories in the data follow the categorization in the CPS. Small firms are firms with less than 10 employees, medium firms are those with between 10 and 99 employees, and large firms are those with over 100 employees. Data is from the 2015 Business Dynamics Statistics. We follow the observed fraction of firms in each category to construct the same groupings in our model.

The probability with which workers can switch employers, s, is chosen to match the observed share of workers that remain in the same size group of employers (small, medium, large) in a given year as in the CPS, which is 78 percent.

We set the level parameter (v_0) to match the average weekly hours of workers. The persistence of the value of leisure shocks (ρ_v) is calibrated to match the auto-correlation of log-hours of workers, while the volatility (σ_v) is calibrated to match the standard deviation in log-hours. Similarly, the persistence of worker efficiency (ρ_x) is chosen to match the auto-correlation of log-wages of workers, while the volatility (σ_x) is set to match the standard deviation in log-wages. In our targets related to the wage variation, we use unconditional data moments since we consider the worker type x as capturing both observed and unobserved heterogeneity in worker efficiency.

The preference shocks for working at different productivity levels of firms generate noise in workers' choices over firms and prevent sorting from being driven entirely by pecuniary returns. As such, changes in the size of these shocks change the steepness of the wage profiles across firm productivity and size groups. Given this, we set σ_ϵ to match the ratio between the average wages in firms with over 100 employees and firms with under 100 employees.

Panel B of Table 4 presents the parameter values that are jointly calibrated and also reported the implied moments of the model against the data. It shows that the model performs well in matching features of the data.

Table 4: Model parameters

Panel A: Outside the model

Parameter	Value	Target
γ	2	Standard
θ	0.67	Labor share of output
ϕ	2	Erosa et al. (2016)
r	0.04	Erosa et al. (2016)
ρ	-0.5	Shao et al. (2022)

Panel B: Calibrated

Parameter	Value	Basis	Data	Model
λ	2.05	Employment share (large firms)	0.66	0.65
β	0.92	Wealth-income ratio	2.5	2.5
s	0.45	Frac. remaining in same firm size	0.79	0.78
v_0	3.63	Average weekly hours	41.7	41.8
σ_v	0.51	SD log hours	0.23	0.23
ρ_v	0.62	Autocorr. log hours	0.67	0.67
σ_x	0.37	SD log wages	0.63	0.62
ρ_x	0.74	Autocorr. log wages	0.70	0.70
σ_ϵ	0.3	Wage gap, large to rest	0.26	0.26

Notes: Panel A reports parameters that are set following the literature. Panel B reports the parameters that are calibrated to match specific data features with the model. The last two columns in Panel B report the data target and model implied value. The employment share of firms with over 100 employees is computed from the 2015 Business Dynamics Statistics. Measures of hours and wages are calculated using the pooled CPS sample. The data target for the wealth-income ratio is from Erosa et al. (2016).

4.2 Model Implications

This section compares the model's implications with the data. We begin by showing that the model generates empirical patterns consistent with the motivating facts detailed in Section 2, and then discuss the relevant features of the model that generate these patterns.

Figure 4 reports the relationship of average wage (solid line) and hours worked (dashed line) by firm size category and shows that the model successfully replicates our first two motivating facts. As the size-wage premium is a target in our calibration, the model exactly matches the increase in wages over firm size; as in the data, wages increase by around 25 percent between the smallest and largest firms.

The model also replicates the (non-targeted) positive relationship between average hours worked and firm size as average hours increase monotonically with firm size. Quantitatively, the model predicts a steeper relationship between hours and size with an increase in average

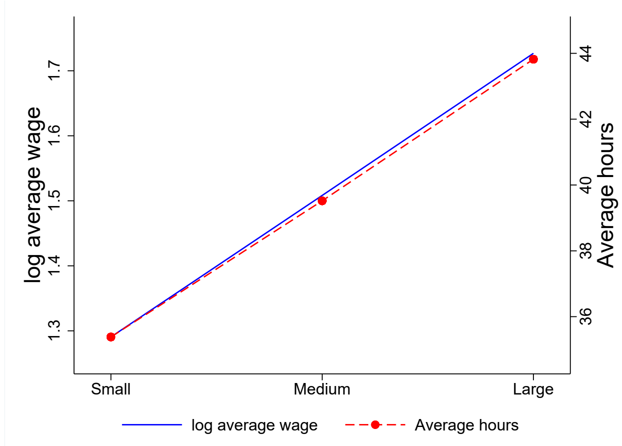

Figure 4: Wages and hours over size in the model

Notes: The figure plots the log average wages (solid line, left axis) and average weekly hours worked (dashed line, right axis) for each size group in the model. Section 4.1 describes the construction of size categories in the model.

hours from the smallest to largest firms of around 8 hours per week in the model compared to a 3 hours increase in the data.

The quantitative match of the size-hours patterns could be improved by introducing additional features into our model. For instance, our model lacks various features that restrict workers' hours, such as kinks in the wage function (modeled exogenously in Bick et al., 2020) or other restrictions imposed from the demand or supply side. Examples of such restrictions include fixed costs of hiring full-time workers or a concentration in preferences for working a given level of hours. The model also predicts a one-to-one mapping between firm productivity and size as we abstract from other sources of firm-level variation such as markups, age, and financial frictions. Such heterogeneity would imply a flatter size-hours relationship by breaking the direct link between firm size and productivity. We preclude such features to keep our analysis relatively parsimonious and transparent.

Figure 5 shows that the model also generates our third and final motivating fact. In particular, for each size category, the model features i) a hump-shaped relationship between hours worked and wages, ii) highest wages close to average hours worked, and iii) an increasing penalty as a worker's hours deviate from the usual hours in the firm. These patterns are not only present in the wage schedules that workers take as given (Panel (a)), but also in the observed relationships between wages and hours (Panel (b)). In addition, the model reconciles our empirical finding that the long-hours penalty is higher in smaller firms, and

(a) Wage schedule

(b) Equilibrium wages

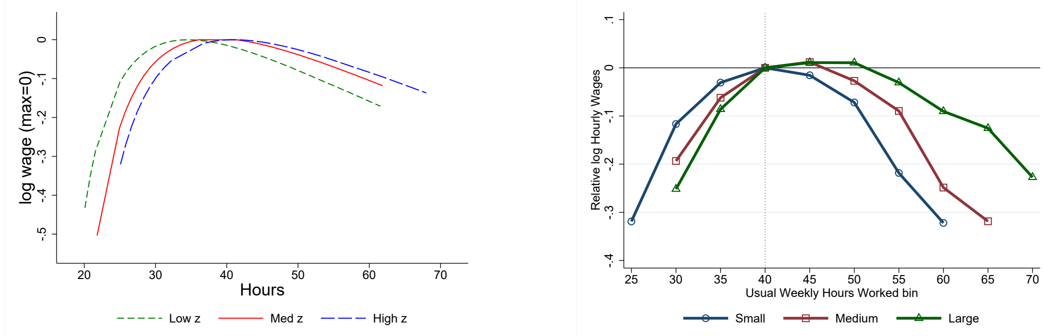

Figure 5: Wage-hours relationship and firm size; wage schedule and equilibrium wages

Notes: Panel (a) reports, for each firm size category, the relative wage schedule by hours worked. To construct the relative wage schedule, we first take an average of the wage function $w_z(l, x)$ across efficiency x, weighted by the measure of each efficiency group. We then plot the logarithm of the average in difference to its maximum level, against working hours. Panel (b) plots, for each size category, the sum of coefficients $(\gamma_h + \theta_{e,h})$ estimated from Equation 2 (only controlling for size group, hours bin and their interactions) using simulated data from the model. This exactly replicates the construction of Figure 3. Section 4.1 describes the construction of size categories in the model.

the short-hours penalty is higher in larger firms.

4.3 Accounting for the Motivating Facts

Having shown that the model successfully replicates our three motivating facts, we explore, in turn, the mechanisms which generate each fact.

Wage differentials between small and large firms. In the absence of heterogeneous preferences for firms, workers are sorted into firms based on wages alone; therefore they would flow into the highest paying firm until wages are equalized across firms. Workers' random preference for firms (i.e. taste shock) introduces noise to their sorting and allows the model's equilibrium to sustain a positive wage gap between high and low productivity firms for any given worker efficiency group. Indeed, the smaller are the taste shocks, the more important wage differentials are to workers' sorting decisions, and the smaller is the size-wage premium. Figure 6 illustrates that the increase in average wage from small to large firms reduces by more than a half when we reduce the size of the taste shocks.[16]

[16]Eliminating heterogeneity in taste shocks entirely is not computationally feasible since, as discussed above, they facilitate the model solution.

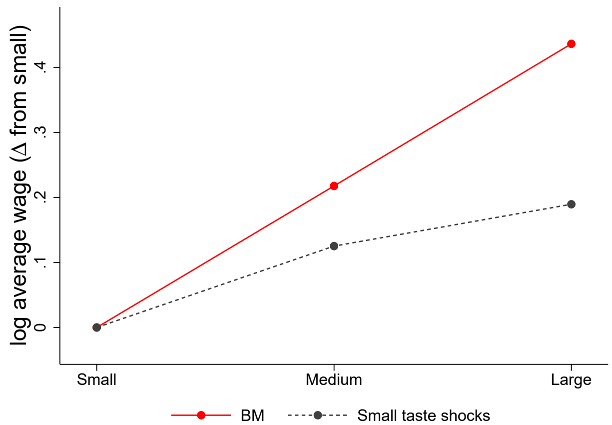

Figure 6: Size-wage premium, role of heterogeneity in tastes and worker efficiency

Notes: The figure log average wages (relative to small firms) for each size group in a model. The solid red line plots the results from the benchmark calibration where $\sigma_\epsilon = 0.30$, the short-dashed grey line is the model with $\sigma_\epsilon = 0.03$, which means the standard deviation of taste shocks that are 10 percent of the benchmark value. We do not recalibrate any other parameter of the model. Section 4.1 describes the construction of size categories in the model.

Increasing hours over firm size. In the model, the presence of a size-wage premium results in longer average hours worked in larger firms. This outcome is, in part, a result of the sorting of workers with different desired hours into different firms. There are two opposing forces that shape the pattern of sorting. The first force pushes workers with longer desired hours to work in larger firms. Specifically, since income is the product of hours and wages, the income gains from working longer are higher in larger (higher wage) firms, so workers with longer desired hours prefer employment in larger firms. The second force pushes in the opposite direction by incentivizing workers with shorter desired hours to work in larger firms. To understand this, notice that employment in smaller (low wage) firms will result in a lower income and hence higher marginal utility for any given level of hours. Marginal utility will be highest for workers with shorter desired hours, and hence these workers value increases in income more than workers with longer desired hours.[17]

In addition to the sorting of workers with different desired hours into firms of different sizes, there is an intensive margin channel that generates longer hours in larger firms. This incentive margin channel encourages all workers to work longer hours in large (high wage) firms regardless of their characteristics. As with sorting, which can be considered an extensive

[17]This is akin to the substitution and income effect following an increase in wages in a standard labor supply decision. However, unlike the standard problem, in our model agents make the decision to choose their workplace and hence, their hourly wages.

margin channel, two opposing forces determine the ultimate effect of the intensive margin channel. On the one hand, higher wages in large firms incentivize workers to work longer hours due to higher income gains. On the other hand, the marginal utility of the same worker in small firms is higher due to a lower income, which encourages them to work longer hours.

The relative magnitude of these two opposing forces depends on how quickly marginal utility decreases as income increases and determines the net effects of the extensive and intensive margins. These, in turn will pin down the overall relationship between of hours worked and firm size. In our calibrated model, marginal utility is sufficiently flat, so the first force dominates. As a result, workers with longer desired hours select into larger firms, and for any given type, workers also work longer hours in large firms. Therefore, average hours increase with firm size in equilibrium, a result of both extensive and intensive margin effects.

(a) The role of sorting

(b) Alternative interest rates

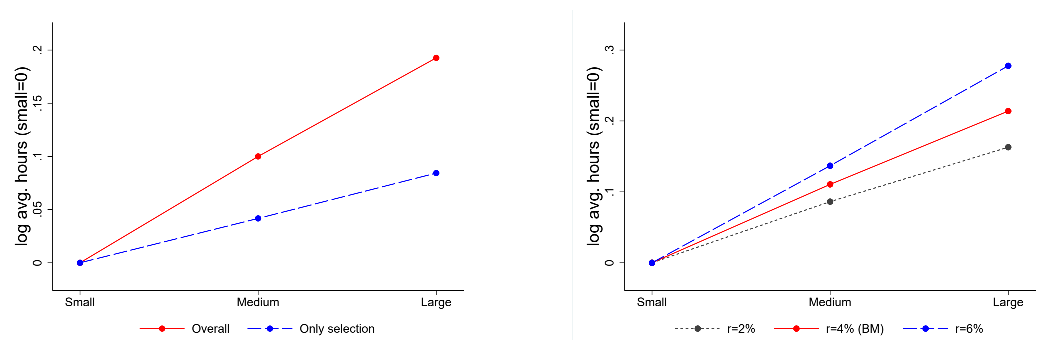

Figure 7: The relationship between hours Worked and firm size

Notes: Panel (a) plots log average hours (in difference to small firms) by firm size in the benchmark model (solid red line), and in an alternative version of the model that only allows selection between firm groups (dashed blue line). To construct average hours in the alternative version, we set hours worked for each worker with the same state variables – skills, value of leisure and wealth – to be equal to the average hours in small firms. We then compute the average hours in each size group according to the resulting distribution of these states across firms. Panel (b) plots the log of average hours (relative to small firms) across firm size for three calibrations of the model. In addition to the benchmark calibration (solid red line), we include two alternatives; a calibration with low interest, $r = 0.02$ (short-dashed gray line) and a calibration with high interest rates $r = 0.06$ (long-dashed blue line).

Figure 7 explores the importance of the mechanisms behind the pattern of increasing hours over firm size. We highlight the role of sorting in Panel (a). In particular, we compare the benchmark model's size-hours relationship (solid red line) to a counterfactual relationship computed by assuming that all workers of the same characteristics work the same hours as their counterparts in small firms (dashed blue line). This captures the size-hours relationship due only to differences in the selection of workers based on their characteristics. Since larger firms attract workers with longer desired hours, the counterfactual also features a positive relationship between firm size and average hours. Panel (a) suggests that sorting accounts for about 40 percent of the positive size-hours relationship in the benchmark model. The remainder, 60 percent, can be attributed to the intensive margin channel, namely that workers of the same characteristics work longer hours in larger firms.

Panel (b) illustrates the role of marginal utility for the positive size-hours relationship. It plots the relationship between firm size and hours for alternative values of the (exogenous) interest rate. Intuitively, higher interest rates raise the wealth of agents, which has a positive impact on consumption levels – lowering marginal utility. Higher interest rates strengthen the first channel influencing sorting highlighted above and generates stronger selection on desired hours. Accordingly, the steepness of the size-hours relationship increases as the interest rates increase from 2 to 6 percent.

Complementarities between workers' hours amplify the positive hour-size relationship. To see this, suppose that, for some reason, a firm features longer average hours than all other firms. Complementarity in hours implies that, compared to another firm, workers with longer desired hours will be penalized less in this firm while workers with shorter desired hours will be penalized more. As a result, long-hour workers will wish to sort into this firm (an extensive margin effect). This sorting will, in turn, result in similar workers of that firm to work longer (an intensive margin effect). Both these extensive and intensive margin effects, driven by complementarities, will amplify the positive relationship between hours and size.

We highlight this amplification by comparing the benchmark calibration to an alternative version of the model that does not feature complementarities. That is, we leave all parameters unchanged but set $\rho = 1$. Panel (a) of Figure 8 compares the size-hours relationship in the two versions of the model. Without complementarities, workers in large firms work around 23% longer than workers in small firms compared to 24% longer in the benchmark calibration with complementarities. Although the change in the hours gap due to complementarities is small, the differences are large when considering the differences between the models in terms of the overall dispersion in hours. Indeed, by effectively removing penalties from working relatively shorter or longer hours, the model without complementarity features significantly higher hours dispersion (0.36 vs. 0.23). Panel (b) normalizes the size-hour relationship by the overall standard deviation in both versions of the model. It shows that the increase in hours from small to large firms in the benchmark corresponds to 0.5 standard deviations instead of 0.3 standard deviations in the model without complementarities. In other words, the relationship between firm size and average hours, in a correlation sense, is much stronger with complementarities.

(a) Average hours

(b) Average hours rel. to overall dispersion

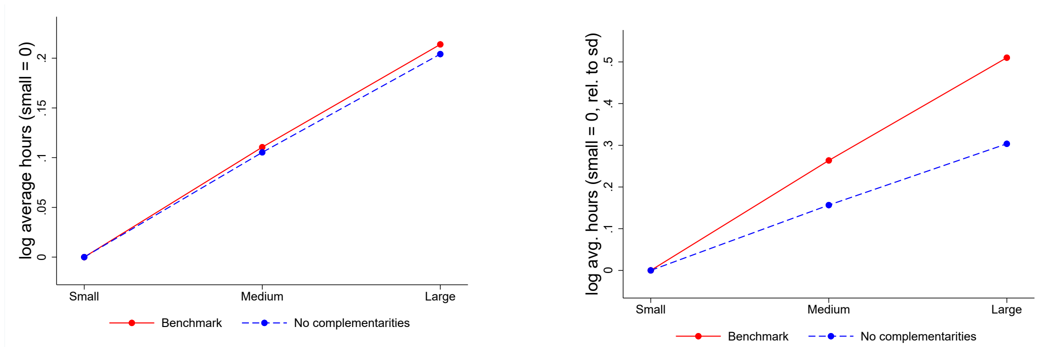

Figure 8: The relationship between hours worked and firm size, role of complementarities

Notes: Panel (a) plots log average hours (in difference to small firms) in the benchmark calibration (solid red line) and an alternative calibration changes only the substitution parameter $\rho = 1$ (dashed blue line). Panel (b) reports the same size-hours relationship but divides average hours (y-axis) by the standard deviation of log hours in each model (i.e. by 0.23 in the benchmark, and 0.36 in the alternative).

Before discussing the models' implications for firm-specific wage schedules, we will briefly remark on the sorting of workers based on their ability x. Some of the same mechanisms that generate sorting based on desired hours also apply to workers sorting based on worker productivity. Indeed, as is the case for desired hours, higher worker efficiency also increases the income gains from employment in larger firms due to the size-wage premium. Accordingly, our model features positive assortative sorting of workers and firms whereby high-efficiency workers tend to sort into high-productivity firms. Since both wages and hours increase in x, worker sorting based on their efficiency strengthens the positive relationship between average wages and average hours over firm size. We find that sorting on ability plays a quantitatively insignificant role in generating the size-wage and size-hours relationships. Indeed, average efficiency increases by 2 percent (or 4 percent of the overall dispersion in efficiency) between small and large firms, which is insufficient to make a quantitative impact on the hours and wage steepness over firm size.

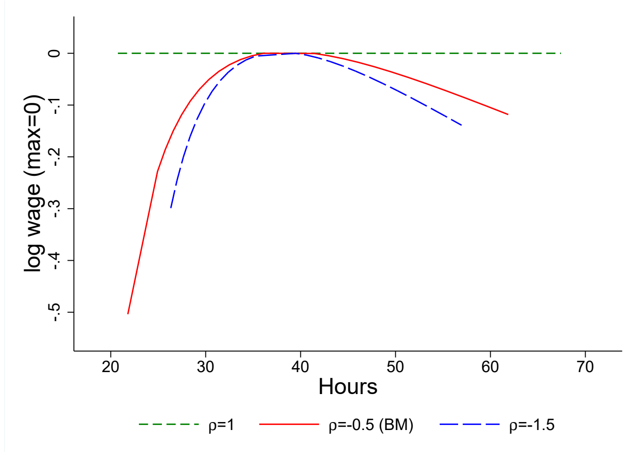

Figure 9: The relationship between wages and hours, role of complementarities

Notes: The figure plots the relative wage schedule by hours worked for different values of ρ, for a medium productivity grid (6th out of 12). To construct the relative wage schedule, we first take an average of the wage function $w_z(l, x)$ across efficiency x, weighted by the measure of each efficiency group. We then plot the logarithm of the average in difference to its maximum level, against working hours. The three lines correspond to the benchmark model ($\rho = -0.5$), as well as the alternatives changing the substitutability parameter to values of 1 and -1.5. In the alternative computations, we do not recalibrate any other parameter. Section 4.1 describes the construction of size categories in the model.

Hump-shaped wage schedule. The model generates a hump-shaped relationship between wages and working hours due to complementarities between workers' hours. Such complementarity maximizes an individual worker's marginal productivity when she works the same hours as the rest of her production unit. Any hours worked above the level of her co-workers' in the worker's marginal productivity diminishing for those extra hours. By the same token, when a worker works shorter hours, she pulls her co-workers' productivity down, which is reflected in her marginal productivity, hence in her wages. Accordingly, there is a penalty for working shorter and longer than the usual hours in the firm. Figure 9 clearly illustrates the role of complementarities, by plotting the wage schedule faced by a worker in a firm of an intermediate productivity as ρ changes. As working hours become less complementary (higher ρ), the link between hours – within a production unit – becomes weaker, and penalties become less severe for both short and long hours.

4.4 Sorting on Hours

Existing theories of sorting between heterogeneous workers and firms emphasize the importance of sorting based on worker skill and firm productivity (see, for example, Eeckhout, 2018). Our model emphasizes a vital role for sorting based not only on ability but also on the desired working hours of workers. In this section, we argue that the model's predictions are consistent with observations and that sorting based on desired hours is an essential yet understudied dimension driving the allocation of workers to firms.

Through the lens of our model, workers sort into firms based on their skills x, idiosyncratic preferences for workplaces, and desired working hours. Desired working hours are ultimately a function of their value of leisure and their marginal utility (which depend on both wealth and x) and play an important role in determining which firms to work for. Indeed, all else equal, the positive relationship between hours and firm size, combined with complementarities in hours, incentivizes those that have longer (shorter) desired working hours to sort into larger (smaller), more (less) productive firms. An empirically testable implication of this mechanism is that workers who work fewer hours than their coworkers will wish to transition to smaller firms where their desired hours are closer to the average coworker. The converse is true for workers that work longer hours than their coworkers – they have an incentive to transition to larger firms where their hours will be similar to those of their coworkers.

We test this implication by comparing worker transitions in the model to those observed in the data. Specifically, we test whether there are systematic differences in the rates at which workers transition across different firm size groups, f, based on their work hours relative to their coworkers. To do this, we take data from the CPS and model simulations for periods t and $t+1$. We classify workers based on whether they work significantly shorter, longer, or similar hours than the average hours worked in their current firm size group. We denote this average as $\bar{h}_{f,t}$. We classify a worker as working shorter (longer) hours if their own hours, h_t, are at least ten percent shorter (longer) than the firm size average $\bar{h}_{f,t}$. We then compute the transition matrices for transitions across firm size categories between t and $t+1$ separately based on whether $h_t < \bar{h}_{f,t}$ or $h_t > \bar{h}_{f,t}$. We construct these transition matrices in the CPS by exploiting the rotating panel nature of the CPS and tracking respondents over 12 months which corresponds to one period in the model.

Table 5 reports the model and data transitions matrices. Recall that our calibration strategy involved pinning down the degree of frictions that limit worker occupational switching, s, to match the share of workers remaining in the same firm size category. This target only

Table 5: Transitions across firm size based on hours worked

	Model			Data		
	Small	Medium	Large	Small	Medium	Large
		$h_t < \bar{h}_{f,t}$			$h_t < \bar{h}_{f,t}$	
Small	0.62	0.13	0.25	0.73	0.17	0.10
Medium	0.07	0.68	0.25	0.15	0.62	0.23
Large	0.06	0.12	0.82	0.04	0.10	0.86
		$h_t > \bar{h}_{f,t}$			$h_t > \bar{h}_{f,t}$	
Small	0.59	0.10	0.32	0.58	0.22	0.19
Medium	0.03	0.65	0.32	0.08	0.61	0.31
Large	0.03	0.09	0.88	0.02	0.09	0.89

Notes: The table reports transition matrices computed in the benchmark model (left column of matrices) and the CPS data (right column of matrices). The first row of matrices reports the transition rates for workers that work at least ten percent shorter hours than the average hours in their firm size group in period t. We denote this average $\bar{h}_{f,t}$ where f indicates firm size group and a workers' own hours at h_t. The second row of matrices show the analogous transition rates for workers that work at least ten percent longer hours than the average hours in their firm size group in period t, that is, $h_t > \bar{h}_{f,t}$. Table A.1 in the Appendix reports the model and data transition matrices for workers whose hours are within 10 percent of their firm size group average. To construct transition matrices in the CPS, we track respondents over a 12-month period and identify the size group in which they are employed in the first t and second $(t+1)$ observations. Workers are grouped based on their firm size group and hours worked in the first period t. Each entry in a transition matrix reports the share of workers that transition from the size group in the matrix row in period t to the size group in the matrix column in period $t+1$.

disciplines the diagonal of the *aggregate* transition matrix in the model and does not provide any guidance on the off-diagonal entries nor on how the transition rates would depend on working hours. Despite this, the model performs well in qualitatively matching observed transition matrices for subgroups with $h_t \lessgtr \bar{h}_{f,t}$.[18]

To begin with, our model would predict that workers working shorter hours ($h_t < \bar{h}_{f,t}$) should transition into smaller firms where average hours are shorter. In the model, around 12% and 6% of employees in large firms transition to medium and small firms, respectively and around 7% of employees in medium sized firms transition to small firms. Sorting based on hours would imply fewer transitions towards smaller firms when $h_t > \bar{h}_{f,t}$. This is indeed the case: Only 9% and 3% of workers in large firms transition to medium and small firms, respectively when they work longer hours than their average coworker. Similarly, only 3% of long-hour workers in medium-sized firms transition to small firms. Instead, the model predicts that when $h_t > \bar{h}_{f,t}$, workers are much more likely to switch to larger firms. For

[18]For the sake of brevity, we omit the transition matrices for workers who worked similar hours to their size group, i.e. $h_t \approx \bar{h}_{f,t}$). These are reported in Appendix Table A.1, which shows that transition rates for this group are roughly in between those of workers with shorter and longer hours.

instance, 32% of all long-hours workers employed in small firms switch to large firms. The analogous transition shares for short-hour workers in small firms is only 25%.

Strikingly, this pattern of likelier switches to smaller firms when $h_t < \bar{h}_{f,t}$ and to larger firms when $h_t > \bar{h}_{f,t}$ is also evident in the data. The matrices on the right-hand side of the table report the observed transition matrices in the CPS. Comparing the model to the data reveals that although the model does not quantitatively match the data, the qualitative pattern consistent with sorting based on hours is evident. For instance, when $h_t < \bar{h}_{f,t}$, 10% and 4% of workers in large firms switch to medium and small firms over a 12 month period and when $h_t > \bar{h}_{f,t}$ these measures are lower at 9% and 2%. The analogous rates in the model at 12%, 6%, 9% and 3%, respectively. Overall, the model predicts more transitions out of small firms. This is most evident when considering the subgroup of workers that work shorter hours with $h_t < \bar{h}_{f,t}$. For example, in the model 25% of short-hour and 32% of long-hour workers in small firms switch to large firms while in the data these measure are 10% and 19%, respectively.

Taken together, Table 5 reveals evidence in the data that is consistent with workers sorting based on hours. In particular, working shorter hours is associated with subsequent movements toward smaller firms, while working longer hours is associated with subsequent movements towards larger firms. This is a crucial prediction of the model and is driven by the interaction of the positive hours-size relationship and complementarities in hours worked. These findings complement our discussion of the positive cross-sectional relationship between firm size and average hours. Indeed, Table 5 shows that a dynamic version of this relationship also holds in the data and that the model performs well in replicating it.

Our findings suggest that hour complementarity in production is an important driver behind the sorting of workers across firms based on their hours worked. We interpret hours complementarity as a generalization of flexibility in workers' ability to set their own working hours. The importance of flexibility in choosing hours is often studied in the literature that aims to understand gender differences in time constraints – hence the demand for flexibility – in shaping the gender-specific sorting decisions towards employment in low-hour occupations (Erosa et al., 2017 and Bento et al., 2021). We contribute to this discussion using a model in which the inflexibility in working hours appears as wage penalties for short- and long-hour workers, an equilibrium outcome that depends on firm characteristics and the rest of the workers in the economy. We highlight the interaction of the inflexibility of working hours and firm heterogeneity in average hours in shaping the overall sorting pattern of workers across firms of different sizes.

5 Aggregate Implications

This section explores the aggregate implications of our theoretical framework, focusing first on implications for wage and income inequality and second on size-dependent firm policies.

5.1 Implications for inequality

Our theoretical framework features the interaction of hours, wages, and firm productivity. Here, we show that this interaction has important implications for inequality in both hourly wages and income – that is, the product of wages and hours. Specifically, we explore how heterogeneity in hours matters not only for income inequality but also wage inequality – a finding not emphasized in existing work. We also highlight the importance of complementarities in influencing inequality.

Wage inequality. The benchmark model exhibits wage inequality due to heterogeneity in three factors: worker skills, firm-level productivity, and working hours. The role of heterogeneity in firm and worker productivity in contributing to inequality is well-studied. Our novel insight in this aspect is to highlight the role of hours worked for wage inequality. To do this, we explore the contribution of each of the above three factors to wage inequality and then compare it to the data.

Table 6 summarizes the results of our analysis. We begin by describing the first row which reports the contribution of worker skills, firm-level productivity, and working hours to wage dispersion in the data. As emphasized in existing work, much of the observed dispersion in wages is driven by unobservable worker characteristics (Abowd et al. 1999). Accordingly, we find that worker skill – proxied by education and years of experience, firm productivity – proxied by firm size, and hours worked only account for 20% of the observed dispersion in wages (first column). To compute this share, we compared the overall standard deviation of observed (log) hourly wages (0.63) to the weighted average of within-group standard deviations which are computed separately for each group where a group is defined by triple of worker skill, firm productivity and hours worked, and the weights are number of observations in a group. We find that the weighted average dispersion is 0.50. Hence, the difference (0.63-0.50, or 20% of 0.63) is *explained* by worker skill, firm productivity and hours worked.

The remaining columns report the contribution of explained dispersion that is due to worker skill, firm productivity, and hours worked, respectively. To construct these contributions, we once again compute weighted averages of within-group standard deviations where a group was defined by worker skill (second column), worker skills and firm productivity (third column), and worker skills, firm productivity and hours worked (last column). We find that 73% of explained dispersion is due to workers' skills while firm size accounts for only 8% of hourly wage dispersion. The remainder, 19%, is due to dispersion in hours worked. Strikingly, dispersion in hours contributes more than twice as much to raising wage dispersion than firm productivity and is around 1/4 as important as worker skills.

The second row of Table 6 conducts a similar breakdown using simulated data from the benchmark model. Notice that there is no notion of unobservable characteristics in the model. So, controlling for x, z, and l fully accounts for all wage dispersion in the model. Heterogeneity in hours worked accounts for 13% of the entire wage dispersion – close to the 19% contribution in the data. The contribution of hours arises due to complementarities which generate the hump-shaped relationship between hours and wages. Indeed, worker wages would be independent of hours worked without complementarity, and the contribution of hours to wage dispersion would be zero. Our model implies a much larger contribution of firm productivity on wage dispersion than in the data (22% vs. 8%) and understates the contribution due to worker skills (65% vs. 73%). This is due to a coarser division of firms in the data, as firm size groups are split into three in the data.

In order to make our analysis of the simulated data more comparable to the observed data, we group firm productivity into the same three size categories as in the data. The decomposition results are reported in the third row of Table 6. Since we are now using a

Table 6: Accounting for dispersion in hourly wages in the data and model

	Share Overall Dispersion Explained by Skills, Firm Productivity & Hours	Contribution due to Skills	Firm Productivity	Hours
Data	0.20	0.73	0.08	0.19
Model (fully observed z)	1.00	0.65	0.22	0.13
Model (grouping z)	0.71	0.83	0.07	0.10
Model w/o compl. (fully observed z)	1.00	0.81	0.19	0.00
Model w/o compl. (grouping z)	0.68	0.89	0.09	0.02

Notes: The table reports the results from an exercise which decomposes wage dispersion in the data (first row) and model simulated data (remaining rows). The first column reports the share of unconditional standard deviation of log hourly wages that is *explained* by worker skill, firm productivity, and hours worked. To compute this share, we compare the overall standard deviation (σ_w) to a weighted average of within-group standard deviations ($\sigma_w^{x,z,l}$) which are computed separately for each group where a group is defined by triple of worker skill (x), firm productivity (z) and hours worked (l), and the weights are number of observations in a group. The share explained is $\frac{\sigma_w - \sigma_w^{x,z,l}}{\sigma_w}$. To proxy for worker skill and firm productivity in the data, we use four education bin and years of experience, and three firm size categories, respectively. The remaining columns report the contribution of skills, firm productivity and hours worked, respectively. These are computed as $\frac{\sigma_w - \sigma_w^{x}}{\sigma_w - \sigma_w^{x,z,l}}$, $\frac{\sigma_w^{x} - \sigma_w^{x,z}}{\sigma_w - \sigma_w^{x,z,l}}$ and $\frac{\sigma_w^{x,z} - \sigma_w^{x,z,l}}{\sigma_w - \sigma_w^{x,z,l}}$, respectively.

coarser division of firm types, the model no longer explains the entirety of all wage dispersion. With this change, the model much more closely matches the data with 7% of explained wage dispersion being due to firm productivity. When grouping firms into bins, worker skills accounts for a larger share of wage dispersion exactly due to the positive correlation between x and z. In the absence of information on the exact firm type z, x plays a larger role in accounting for wage dispersion. It is also worth noting that the share explained by working hours decreases when we categorize firms instead of using actual firm productivity. This is because individual hours' effects on wages strongly depend on the usual hours in firms, which in turn depends on firm productivity. By using a coarser firm categorization, we lose the explanatory power of hours on wages.[19]

In the data, heterogeneity in hours accounts for almost one-fifth of explained dispersion in wages. Through the lens of the model, such a contribution to dispersion in wages is driven

[19]This same logic suggests that using finer firm groupings in the data should increase the observed contribution of hours in wage dispersion. We test this by using the full range of firm size categories reported in the CPS and find that using finer size categories results in the contribution of hours to wage dispersion to increases by 3 percentage points to 22%. Recall, that the reporting of firm size categories in the CPS has varied over time and it is for clarity and consistency that we focused on reporting three categories of firm size.

by the presence of complementarity in workers' hours. To illustrate this, the last two rows of Table 6 repeat the decomposition analysis in a version of the model where there are no complementarities. When we assume that firm type is fully observed, hours worked play no role in generating wage dispersion. This is to be expected since, without complementarities, wages within firms are independent of working hours. That is, there is no hump-shaped relationship between wages and hours. The conclusion is similar when we group firm types. Hours only account for 2% of explained wage dispersion, which captures the correlation between hours and firm productivity z, while the bulk (89%) of dispersion is due to worker skills.

Taken together, the analysis reported in Table 6 highlights a number of important findings. First, in the data, variation in hours is a significant contributor to wage inequality. Second, our model qualitatively matches the data with respect to decomposing the contribution of wage dispersion. Finally, we show that complementarities in hours are essential for heterogeneity in hours to contribute to wage dispersion. Next, we explore the models' implications for income inequality – that is, the product of wages and hours.

Income inequality. Variation in hours naturally generates variation in income, not only mechanically but also due to the correlation between hours and wages. While hours inequality contributes to income inequality positively, the correlation between wages and hours can mitigate or amplify the overall income inequality. For example, if high-wage workers work more hours than workers earning lower wages – a positive correlation between wages and hours – then overall income inequality would be greater. This can be seen clearly through a simple variance decomposition of (log) income,

$$\begin{aligned}\text{Var(income)} &= \text{Var(wage)} + \text{Var(hours)} + 2 \times \text{Cov(wage,hours)} \\ &= \text{Var(wage)} + \text{Var(hours)} + 2 \times \text{Corr(wage,hours)} \times \sqrt{\text{Var(wage)} \times \text{Var(hours)}}\end{aligned} \tag{9}$$

Motivated by this, we explore the drivers of variance in income, emphasizing the endogenous correlation between hours and wages. Table 7 reports each component on the right hand side of (9) for both the data and model (first two rows). The second and the third columns of the table show that hours dispersion and the covariance of hours with wages account for a significant portion of overall income dispersion (around 20 percent in total) in both the data and model. Recall, in our calibration, we only targeted dispersion in hours and wages and did not target the covariance between hours and wages nor the variance in income. Despite this, the model gives a decent fit to both the dispersion in income and, importantly, the covariance between hours and wages.

Table 7: Decomposing the variance in income, the role of hours and wages

	Var(income)	Var(wage)	Var(hours)	$2 \times$ Cov(wage,hours)	Corr(wage,hours)
Data	0.50	0.40	0.05	0.05	0.18
Model	0.51	0.38	0.05	0.08	0.28
Model w/o compl.	0.59	0.36	0.13	0.10	0.22

Notes: The table reports the results from a variance decomposition of log income following (9) in the data (first row), benchmark model (second row) and a version of the model without complementarities (last row). The first column reports the total variance in log income while the next three columns report each of the three components of the right hand side of (9). The last column reports the correlation between hours and wages.

As with wage inequality, complementarities play an important role in influencing income inequality. In fact, there are three main channels through which complementarities alter the right hand side of (9). First, complementarities generate wage dispersion for a given worker-firm combination. Indeed, if hours were perfectly substitutable across workers, wages would be identical for all workers in a firm (of a given skill x) regardless of their hours worked. Thus complementarities raise the variance in wages.

Second, complementarities introduce an incentive for workers to work similar hours to their coworkers (through hump-shape wage profiles). This compresses the distribution of hours worked in an economy and hence complementarities also serve to reduce the variance in hours pushing towards lowering income inequality.

Third, complementarities strengthen the positive relationship between hours and firm size. We have highlighted this before in Figure 8 where the change in hours from small to large firms, relative to the dispersion of hours in a given model, is weaker without complementarities. Accordingly, complementarities also raise the *correlation* between hours and wages.

To see the overall impact of these three channels on income inequality, the last row of Table 7 reports the terms in (9) in a version of the model without complementarities. First, removing complementarities raises income dispersion in our model from 0.51 to 0.59. Decomposing this dispersion reveals that the increase in income dispersion is driven primarily by an increase in hours dispersion. That is, the second channel discussed above dominates other the other two. The compression of the hours distribution due to complementarities is quite stark, with the hours dispersion in our benchmark model being around 40 percent of the counterpart model without complementarities.

The table also highlights the aforementioned channels through which complementarities contribute positively to income dispersion, even though these effects are ultimately offset with compression in hours distribution. First, the model with complementarities features higher wage dispersion. Second, the correlation between hours and wages is stronger in hour baseline. However, the latter effect does not translate into an increase in income inequality because the reduced hours dispersion with complementarities shrink the magnitude of the covariance between hours and wages. In fact, the covariance is lower in our baseline than in the counterfactual without complementarities. In sum, the reduction in hours dispersion is strong enough to cancel the opposing positive effects of complementarities in income inequality.

Consistent with existing empirical literature, the model predicts that heterogeneity in hours is an essential contributor to income inequality (see for example, Blau and Kahn, 2011 and Checchi et al., 2016). Since a role for hours arises endogenously, our theoretical framework provides insights into the mechanisms by which hours impact income inequality. Indeed, we find that complementarity in working hours plays a crucial role by compressing the distribution of hours while raising wage inequality and driving more pronounced sorting on hours across firms.

5.2 Implications for size-dependent firm policies

Size-dependent firm policies and regulations are standard features across economies. Such policies include the provision of subsidies to smaller firms or more onerous regulation for larger firms. These have been shown to have a significant impact on output and welfare. For instance, a 'progressive' sales tax that increases with the productivity or size of firms would make the effective productivity of firms more equal and force higher productivity firms to hire fewer workers. In most models, this would lead to output losses and reduce welfare (see, for example, Guner et al. 2008, Guner et al. 2018, Gourio and Roys 2014, and Bento and Restuccia 2017).

In this section, we highlight a novel channel through which these policies may impact welfare through the lens of our model. Complementarities in hours worked imply penalties for working different hours from the usual in the firm. To the extent that workers with high and low propensities to work cannot coordinate to perfectly sort into different firms, they suffer these penalties because they are typically far from their co-workers' hours. Sorting of long-hour workers into larger firms mitigates this problem by facilitating this separation. It brings a higher discrepancy among firms in terms of their usual hours, and widens the menu of firms that workers can self-select into. A policy that makes the net productivity of firms more similar will make wages across firms more similar. Hence, average hours across firms will also be more similar, which shrinks the menu of average hours available to workers.

We highlight this mechanism with a simple set of policy exercises in which we apply productivity specific sales taxes on firms given by a rate, $\tau(z) = 1 - \tau_0 \times z^{-\tau_1}$. This form of sale tax is similar to the size-dependent distortions in Guner et al. (2018). Firms face the same average sales tax rate τ_0 when $\tau_1 = 0$. When $\tau_1 > 0$, larger firms – firms with a higher z – face a higher tax rate than small firms. We simulate counterfactual economies with τ_1 set at 10 and 20 percent, while we let τ_0 adjust to maintain the net revenues from a taxation scheme at zero. We study the consequences of these changes for output, income inequality, and welfare. In doing so, we focus on steady state comparisons.[20] We study the implications of these exercises for our benchmark, as well as for the alternative calibration without complementarities ($\rho = 1$).[21]

[20]For welfare, we use the value functions of individuals right before drawing the shocks for a given period, i.e. knowing their efficiency x, value of leisure v, and last period's employer's productivity z_j, but before drawing the new shocks x', v', ϵ', and before knowing whether or not she can choose the new productivity level to work for.

[21]The reason behind our functional form is that it only requires us to set one parameter, τ_1, rather than defining an arbitrary target group of firms. The reason for adjusting τ_0 to have the net revenues from taxation to be zero is to minimize the welfare effects from discarding resources. In terms of the values for the levels of distortions, the degrees of progressivity are within the range experimented by Bento and Restuccia (2017) in the context of development. The output losses implied in our framework are similar to those found in that paper.

Table 8: Distortions experiment, changes relative to baseline

Distortions	Tot. output (%)	Income dispersion (p.p.)	Welfare CE, %	Welfare Frac. better off
		Benchmark		
Mild	-1.3	-2.5	0.08	0.58
Strong	-3.0	-3.9	-0.53	0.22
		No complementarities		
Mild	-1.3	-2.6	0.26	0.71
Strong	-3.0	-4.0	-0.40	0.30

Notes: We conduct two policy experiments for each of the two alternative calibrations: the benchmark model and the alternative model without complementarities (i.e. $\rho = 1$), that features otherwise the same calibration as in the benchmark. The experiments implement a TFP-specific distortion on firms as represented with $\tau(z) = 1 - \tau_0 \times z^{-\tau_1}$, with mild ($\tau_1 = 0.1$) and strong ($\tau_1 = 0.2$) distortions. The table gives changes relative to the corresponding baseline. Second column gives the percentage change in steady state total output and third column gives the percentage point change in the steady state income dispersion. Third column gives the median welfare gain from moving to the post-experiment steady state in consumption equivalence (%) terms. Final column gives the fraction of individuals that prefer the post-experiment steady state to the baseline. We describe the derivation of the gains in Appendix C.2.

We summarize our results in Table 8. Our main finding is that the welfare losses due to distortions are higher (or gains are smaller) in our benchmark than in the alternative without complementarities. The aforementioned intuition is behind this result. By reducing the firm-level dispersion in net productivity ($(1 - \tau(z))z$), the distortions hinder workers' self-selection into firms according to their desired worker hours. This comparison is similar both for the mild distortions ($\tau_1 = 0.1$), and for the more severe distortions ($\tau_1 = 0.2$).[22]

Importantly, these differences in the welfare implications are not due to the discrepancies in the output or income inequality implications in the two models. In fact, the distortions lead to similar output losses across models, with 1.3 percent loss in the steady state output for the mild distortions, and 3 percent loss with severe distortions. Since firm-level heterogeneity in net productivity contributes to the dispersion in wages, both models predict reductions in wage and earnings dispersion due to distortions, and these are of similar degrees in the two models. All in all, the effects of distortions on output levels and income inequality alone would not have implied differences in the welfare implications between our benchmark and in the alternative.

Next, we discuss in more detail the mechanisms through which the complementarities

[22]Our model features taste shocks that directly affect the values of workers. We keep the distribution of these shocks identical in the model with or without distortions. Volatility in these shocks blurs welfare differences across models due to changes in consumption and labor supply. For this reason, in addition to a consumption equivalence measure, we also present the welfare results using a measure of winners and losers from policies – a measure less sensitive to taste shocks.

(a) Size and wages (b) Size and hours

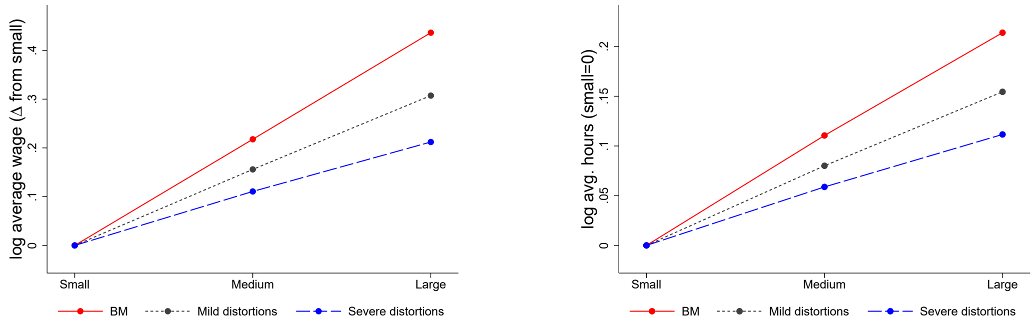

Figure 10: Size-specific effects of a distortionary sales tax

Notes: The left figure plots the log average wages (relative to small firms) for each size group in the model. The right figure does the same for log average hours across firm size groups. Section 4.1 describes the construction of size categories in the model. The three lines correspond to the benchmark model, and the counterfactual experiments of TFP-specific distortion on firms as represented with $\tau(z) = 1 - \tau_0 \times z^{-\tau_1}$, with mild ($\tau_1 = 0.1$) and severe ($\tau_1 = 0.2$) distortions.

generate additional welfare losses. Panel (a) of Figure 10 shows that the wage premium of working in large (and productive) firms shrinks with the degree of distortions while Panel (b) shows that this translates into flatter patterns of average hours across size groups of firms. The hours being more uniform across firms dampens the sorting of workers into firms in the hours dimension. For instance, small firms serve less as a platform to accommodate the short hours needs of workers with high values of leisure.[23]

(a) Low productivity firm (b) High productivity firm

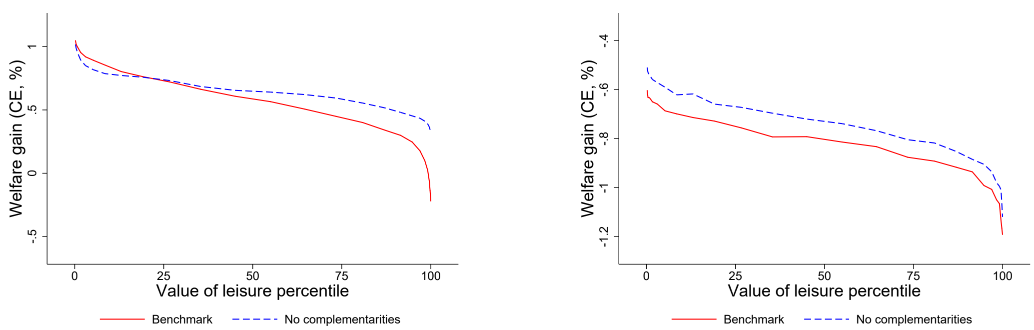

Figure 11: Welfare effects of a distortionary sales tax

Notes: The x-axis in the figures correspond to percentiles of value of leisure in the population-wide ergodic distribution. For a given firm productivity level, we first compute the welfare gains (consumption equivalence in percentage) of each individual from switching to a world (steady state) with the distortions before the beginning of the period. We show the gains in percentage points difference from the median value of leisure worker. Panel (a) uses a low productivity firm (second lowest) and Panel (b) uses a high productivity firm (second highest). We use the experiment with severe distortions for the figures. Welfare gains are computed as consumption equivalence (%) terms. We describe the derivation of the gains in Appendix C.2.

[23]In appendix Figure A.2, we show that the wage penalties for short-hours in small firms increase with higher distortions.

The relevance of the above mechanism for the welfare gains differs within our model population. Figure 11 shows the welfare gains/losses due to mild distortions for different value of leisure groups, separately for workers who were employed in the previous period in a low productivity firm (Panel (a)), and a high productivity firm (Panel (b)). Some welfare implications are common between the models with and without complementarities. Since the distortions hurt particularly high productivity firms, workers associated to these firms typically lose more from the distortions, as illustrated by the levels of the y-axis in the two panels. Moreover the higher value of leisure implies larger losses due to distortions, with and without the complementarities, because of the output reductions hurt particularly more the low consumption workers, which among other traits, tend to have higher values of leisure. Nevertheless, the complementarities between hours alter the slope of the gains against the value of leisure. In low productivity firms, the change in gains from low to high value of leisure is much more significant with complementarities. This is because the income losses are much stronger for the short hour workers in short hour (small) firms, as distortions make these workers more likely to co-work with long hour workers. Without complementarities, the difference between the highest and lowest value of leisure workers' welfare gains is about 0.5 percentage points. In our benchmark, this gap widens to 1.2 percentage points. Panel (b) exhibits opposite patterns for high productivity firms, with the gap in gains being narrower in our benchmark than without complementarities. This is because in our benchmark, long hour workers will have to work with more short hour workers after distortions, leading to additional welfare losses for low value of leisure workers in high productivity firms.

6 Conclusion

This paper studies the relationship between hours, wages and firm-level heterogeneity – specifically firm size. Using micro-data from the US, we document that workers' average wages and average hours increase with firm size, and, novel to the literature, that wage penalties for long (short) hours are larger in smaller (larger) firms.

Motivated by this evidence, we develop a general equilibrium model of heterogeneous firms and workers. Our framework generates a size-wage premium through heterogeneity in workers' preferences for the workplace. The size-wage premium leads workers willing to work longer hours to endogenously sort into larger (more productive) firms, as well as making similar workers work longer hours in larger firms. The existence of complementarities between workers' hours combined with the longer hours larger firms result in less severe long-hour wage penalties and more severe short-hour wage penalties in larger firms – as observed in the data.

We use our model to study the aggregate implications of the interaction of hours, wages, and firm size. We argue that this interaction has important consequences for earnings inequality, and firm policy. We show that, consistent with the data, the model suggests a significant role for heterogeneity in hours in driving wage and income dispersion. We also study the importance of complementarities in working hours for inequality. Finally, we highlight a novel channel through which size-dependent firm policies generate welfare losses. In particular, we argue that these policies shrink the "menu" of options available for heterogeneous workers with endogenous labor supply.

Taken together, this paper highlights the endogenous interactions between hours, wages, and firm-level heterogeneity and argues that these interactions are important in influencing aggregate outcomes.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–334.

Agell, J. (2004): "Why are Small Firms Different? Managers' Views," *Scandinavian Journal of Economics*, 106, 437–452.

Altonji, J. G. and C. H. Paxson (1988): "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-Offs," *Journal of Labor Economics*, 6, 254–76.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade shocks and labor adjustment: A structural empirical approach," *American economic review*, 100, 1008–45.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013): "Cross-country differences in productivity: The role of allocation and selection," *American Economic Review*, 103, 305–34.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016): "It's Where You Work:

Increases in the Dispersion of Earnings across Establishments and Individuals in the United States," *Journal of Labor Economics*, 34, S67–S97.

Battisti, M., R. Michaels, and C. Park (2021): "Labor supply within the firm," *Working Paper*.

Bento, P. and D. Restuccia (2017): "Misallocation, establishment size, and productivity," *American Economic Journal: Macroeconomics*, 9, 267–303.

Bento, P., L. Shao, and F. Sohail (2021): "Gender Gaps in Time Use and Entrepreneurship," Working paper, Working Paper.

Bick, A., A. Blandin, and R. Rogerson (2020): "Hours and wages," Tech. rep., National Bureau of Economic Research.

Blau, F. D. and L. M. Kahn (2011): "Inequality and Earnings Distribution," ISBN: 9780199606061.

Brown, C. and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.

Caliendo, L., M. Dvorkin, and F. Parro (2019): "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 87, 741–835.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 36, S13–S70.

Chang, Y., S.-B. Kim, K. Kwon, and R. Rogerson (2020): "Cross-sectional and aggregate labor supply," *European Economic Review*, 126, 103457.

Checchi, D., C. García-Peñalosa, and L. Vivian (2016): "Are changes in the dispersion of hours worked a cause of increased earnings inequality?" *IZA Journal of European Labor Studies*, 5, 15.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The Quarterly Journal of Economics*, 126, 749–804.

Cubas, G., C. Juhn, and P. Silos (2019): "Coordinated Work Schedules and the Gender Wage Gap," NBER Working Papers 26548, National Bureau of Economic Research, Inc.

Eeckhout, J. (2018): "Sorting in the labor market," *Annual Review of Economics*, 10, 1–29.

Erosa, A., L. Fuster, and G. Kambourov (2016): "Towards a Micro-Founded Theory of Aggregate Labour Supply," *Review of Economic Studies*, 83, 1001–1039.

Erosa, A., L. Fuster, G. Kambourov, and R. Rogerson (2017): "Hours, occupations, and gender differences in labor market outcomes," Tech. rep., National Bureau of Economic Research.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. R. Warren (2020): "Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]," *Minneapolis, MN: IPUMS*.

Gourio, F. and N. Roys (2014): "Size-dependent regulations, firm size distribution, and reallocation," *Quantitative Economics*, 5, 377–416.

Guner, N., A. Parkhomenko, and G. Ventura (2018): "Managers and productivity differences," *Review of Economic Dynamics*, 29, 256–282.

Guner, N., G. Ventura, and Y. Xu (2008): "Macroeconomic implications of size-dependent policies," *Review of Economic Dynamics*, 11, 721–744.

Headd, B. (2000): "The characteristics of small-business employees," *Monthly Labor Review*.

Heathcote, J., K. Storesletten, and G. L. Violante (2014): "Consumption and labor supply with partial insurance: An analytical framework," *American Economic Review*, 104, 2075–2126.

Idson, T. L. (1990): "Establishment size, job satisfaction and thestructure of work," *Applied Economics*, 22, 1007–1018.

Kaplan, G. and G. L. Violante (2010): "How much consumption insurance beyond self-insurance?" *American Economic Journal: Macroeconomics*, 2, 53–87.

Labanca, C. and D. Pozzoli (2021): "Constraints on hours within the firm," *Journal of Labor Economics*.

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 112, 169–212.

Leung, D., C. Meh, and Y. Terajima (2008): "Firm size and productivity," Tech. rep., Bank of Canada.

McFadden, D. L. (1978): "Modeling the Choice of Residential Location," *Spatial Interaction Theory and Residential Location*.

Melo, P. C., D. J. Graham, and R. B. Noland (2009): "A meta-analysis of estimates of urban agglomeration economies," *Regional Science and Urban Economics*, 39, 332–342.

Montgomery, M. (1988): "On the Determinants of Employer Demand for Part-Time Workers," *The Review of Economics and Statistics*, 70, 112–17.

Morchio, I. and C. Moser (2018): "The Gender Pay Gap: Micro Sources and Macro Consequences," MPRA Paper 99276, University Library of Munich, Germany.

Oi, W. Y. (1974): "On the Economics of Industrial Safety," *Law and Contemporary Problems*, 38, 669–699.

Oi, W. Y. and T. L. Idson (1999): "Firm size and wages," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 3 of *Handbook of Labor Economics*, chap. 33, 2165–2214.

Rosen, S. (1986): "The theory of equalizing differences," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard, Elsevier, vol. 1 of *Handbook of Labor Economics*, chap. 12, 641–692.

Shao, L., F. Sohail, and E. Yurdagul (2022): "Are Working Hours Complements in Production?" CEPR Discussion Papers 17311, C.E.P.R. Discussion Papers.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. von Wachter (2019): "Firming Up Inequality," *The Quarterly Journal of Economics*, 134, 1–50.

Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20, 177–181.

Wolpin, K. I. (1984): "An estimable dynamic stochastic model of fertility and child mortality," *Journal of Political Economy*, 92, 852–874.

Yurdagul, E. (2017): "Production Complementarities and Flexibility in a Model of Entrepreneurship," *Journal of Monetary Economics*, 86, 36 – 51.

Online Appendix for:

Labor Supply and Firm Size

Lin Shao Faisal Sohail Emircan Yurdagul

A Additional Figures and Tables

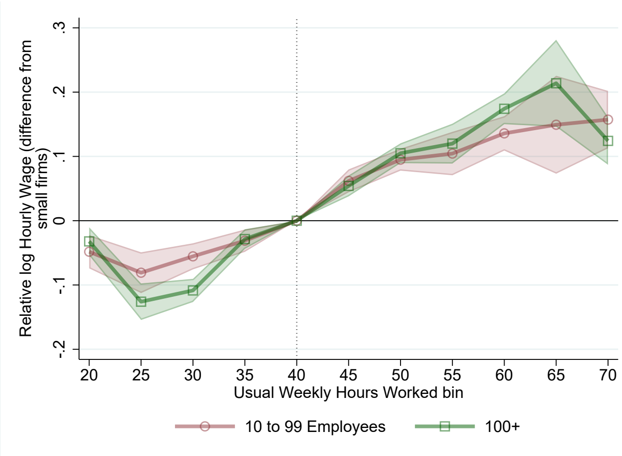

Figure A.1: Wage penalty relative to small firms

Notes: The figure reports the coefficient $\theta_{e,h}$ estimated from Equation 2 where the reference group for hours worked is workers that work $40-44$ hours. The reference group for size is firms with under 10 employees. The shaded regions are the 95% confidence intervals. Data from the pooled CPS sample.

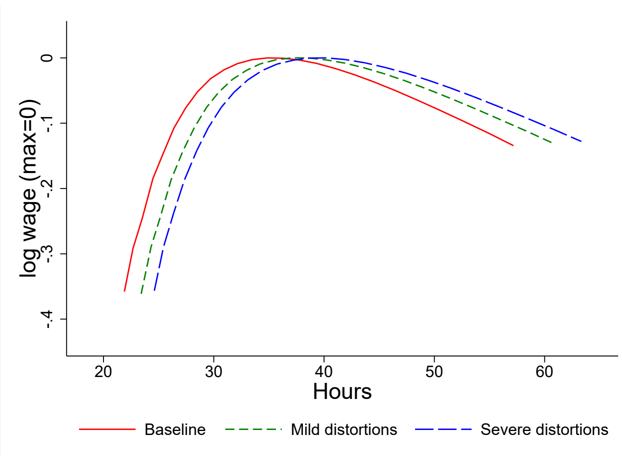

Figure A.2: Wage-hours relationship, role of distortions for small firms

Notes: The figure plots the relative wage schedule by hours worked in a low productivity firm (2nd TFP grid out of 12) in the baseline and in the two distortion experiments. To construct the relative wage schedule, we first take an average of the wage function $w_z(l,x)$ across efficiency x, weighted by the measure of each efficiency group. We then plot the logarithm of the average in difference to its maximum level, against working hours. The three lines correspond to the experiments of TFP-specific distortion on firms as represented with $\tau(z) = 1 - \tau_0 \times z^{-\tau_1}$, with mild ($\tau_1 = 0.1$) and severe ($\tau_1 = 0.2$) distortions.

Table A.1: Transitions across firm size based on hours worked for workers working similar to average hours

	Model			Data		
	Small	Medium	Large	Small	Medium	Large
	$h_t \approx \bar{h}_{f,t}$			$h_t \approx \bar{h}_{f,t}$		
Small	0.60	0.12	0.28	0.56	0.25	0.18
Medium	0.05	0.66	0.29	0.09	0.62	0.29
Large	0.04	0.10	0.85	0.02	0.10	0.87

Notes: The table reports two transition matrices computed in the benchmark model (left matrix) and the CPS data (right matrix). The matrices reports the transition rates for workers that work within least ten percent of the average hours in their firm size group in period t. We denote this average $\bar{h}_{f,t}$ where f indicates firm size group and a workers' own hours at h_t. Table 5 in the main text reports the model and data transition matrices for workers who work 10 percent shorter or longer than their firm size group average. To construct transition matrices in the CPS, we track respondents over a 12-month period and identify the size group in which they are employed in the first t and second $(t+1)$ observations. Workers are grouped based on their firm size group and hours worked in the first period t. Each entry in a transition matrix reports the share of workers that transition from the size group in the matrix row in period t to the size group in the matrix column in period $t+1$.

B Data Appendix

In this appendix, we present and discuss additional empirical results for each of the three primary data sets used in our analysis.

B.1 CPS

Hourly and salaried workers As highlighted in Bick et al. (2020), workers that are paid by the hour experience a relatively stable penalty when working over 60 weekly hours. In contrast, salaried workers experience much larger penalties when working long hours above 60. Given this, our empirical finding that the long (and short) hour penalty varies with firm size could follow simply due to differences in the compositions of workers across firms. For example, if larger firms feature a higher share of hourly workers working longer hours than smaller firms then this could generate the relatively flatter long hours penalty. Figure B.3 tests whether this is the case by plotting the share of workers that are paid hourly by firm size and usual hours worked bins. The figure shows that the share of hourly workers declines as hours worked increase across all firm sizes. Further, the share of hourly workers is relatively similar across firm size bins for usual hours above 40, suggesting that the composition of workers is likely not the primary driver of the flatter long-hours penalty in larger firms.

Having said this, larger firms feature a relatively higher share of short-hour, hourly workers than smaller firms. To concretely test whether differences in composition drive the differences in the short and long hour penalties by firm size, we re-estimate the regression in Equation 2 while also including an indicator for whether workers are salaried or paid by the

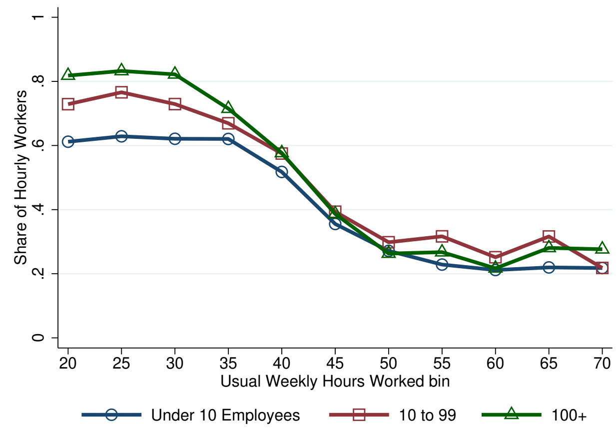

Figure B.3: Share of hourly workers by firm size and hours worked

Notes: The figure plots the share of workers that are paid by the hour, by firm size and usual hours worked. Data is from the outgoing rotation group (ORG) sub-sample in the pooled CPS sample. The ORG sub-sample makes up around 25% of the pooled CPS sample and contains information on whether respondents are paid by the hour.

hour. Figure B.4 reports the sum $(\gamma_h + \theta_{e,h})$ (Panel (a)) and the coefficient $\theta_{e,h}$ (Panel (b)) as estimated from this regression. Due to the smaller sample size when restricting attention to respondents with information on hourly or salaried status, we group usual hours worked into 10-hour bins. The reference group for usual hours worked in the regression is workers that work $40 - 49$ hours.

Figure B.4: Wage profiles by firm size and hours worked, controlling for hourly workers

Notes: The figure reports the coefficient $(\gamma_h + \theta_{e,h})$ in Panel (a) and $\theta_{e,h}$ in Panel (b) as estimated from Equation 2 with an additional indicator variable for whether or not a worker is paid by the hour. The reference group for usual hours worked in the regression is workers that work $40 - 49$ hours. The reference group for size is firms with under 10 employees. The shaded regions are the 95% confidence intervals. Data is from the outgoing rotation group sub-sample in the pooled CPS sample.

Panel (a) shows that the hump-shaped nature of the wage-hours profile remains unchanged when controlling for the salaried status of workers. Also persisting are apparent differences in the wage penalties between the smallest firm size categories and larger firms. This can be seen more clearly in Panel (b), which shows that medium and large firms exhibit more severe short-hours and less severe long-hour penalties compared to small firms.

However, the difference in penalties between medium and large firms becomes much smaller when controlling for whether workers are paid by the hour – particularly for low levels of usual hours worked.

Taken together, this evidence suggests that differences in the composition of workers are not likely drivers of the differences in wage penalties observed across firm size bins.

Controlling for occupations Our primary empirical analysis does not include controls for worker occupations. We make this choice to capture the idea that production involves the interaction of workers employed in different types of occupations. Specifically, the complementarity in hours may be particularly salient between occupations rather than within occupations. Having said this, here we show that our findings from Section 2 are robust to controling for occupation. Specifically, we estimate Equations (1) and (2) while also including an additional regressor that includes dummies for 3-digit occupations as recorded in the IPUMS variable `occ90ly`.

The first and second columns of Table B.2 reports the coefficient β_f from estimating, respectively, a version of Equations (1) and (2) which controls for occupations. Controlling for occupations has little impact on the positive relationship between firm size and hours or firmsize and wages.

Table B.2: The size-wage premium

	log hours	log wages
10 to 99 Employees	0.049***	0.085***
	(0.001)	(0.003)
100+ Employees	0.068***	0.200***
	(0.001)	(0.003)
Year, State FE	Y	Y
Demographic Controls	Y	Y
2-digit Industry FE	Y	Y
Occupation FE	Y	Y
N	1,000,820	1,000,820
R^2	0.175	0.481

Notes: The first and second columns of the table report the coefficient β_f from estimating Equations 1 and 2, respectively, while also including controls for occupations. The reference size category is the smallest size firms. That is, firms with under 10 employees. The reference hours bin is $40 - 44$ hours. Data is from the pooled CPS sample. Standard errors are reported in parentheses. *** indicates statistical significance at the 1% level.

Figure B.5 plots the relationship between hours and wages by firm size in the CPS as estimated from Equation (2) while also controlling for occupation. In particular, Panel (a) reports the sum of the coefficients γ_h and $\theta_{f,h}$ which captures the wage penalty of working outside of the 40-45 hours bin by firm size. Panel (b) reports the coefficient $\theta_{f,h}$ estimated from the same regression. The figure shows that controlling for occupation does not significantly alter the wage-hours relationship across firms. Panel (b) shows that the difference in relative penalties continue to be statistically significant.

(a) Wage-hours relationship by firm size (b) Wage penalties relative to small firms

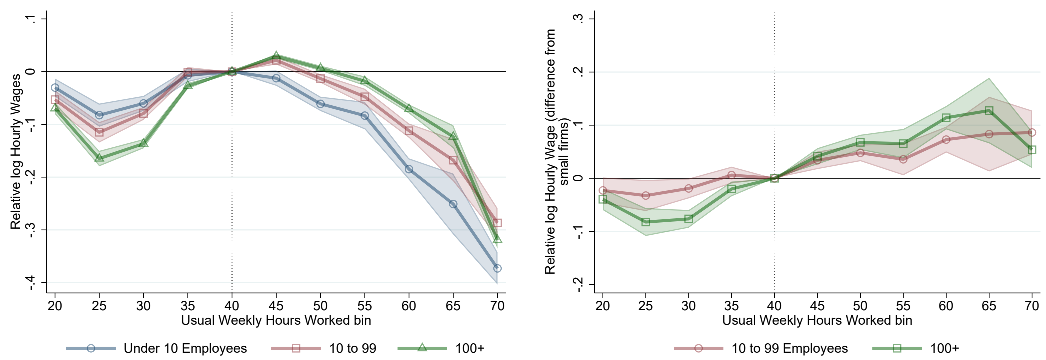

Figure B.5: The relationship between wages and hours, controlling for occupations

Note: Panel (a) reports the the sum of coefficients $(\gamma_h + \theta_{f,h})$ estimated from a version of Equation (2) which also includes controls for occupation. The reference group for usual hours worked in the regression is workers that work $40 - 44.9$ hours. The reference group for size is the smallest size category. Panel (b) reports the coefficient $\theta_{f,h}$ estimated from the same regression. The shaded regions are the 95% confidence intervals. Data is from the pooled CPS sample.

B.2 LFS

Our primary empirical analysis utitlizes data from the US. In this appendix, we document our three motivating facts by firm size using data from the Canadian Labor Force Surveys (LFS) between 1998 and 2018. Similar to the CPS, the LFS is a nationally representative survey containing detailed information on respondents' economic activity for the month they are interviewed such hourly earnings, usual weekly hours worked, and firm size. Firm size is recorded in one of four bins and for clarity, we combine the larger two size bins into one and report size in three categories; i) small (under 20 employees), ii) medium (between 20 and 100 employees), and iii) large firms (over 100 employees). Our sample starts in 1998 as this is the first year that information on establishment size is available in the LFS. Our treatment of the LFS data remains identical to that of the CPS. In particular, we restrict attention to respondents aged 25 and 64 who worked for a single private employer during the reference month. We exclude workers who usually work fewer than 10 hours per week and those that earned less than half the minimum wage.[1] Since 1997, the LFS has also reported establishment size and we conduct our empirical analysis below at both the firm and establishment level.

<u>Fact 1</u> Average hours increase with size.

We begin by showing that workers in larger firms and establishments work longer hours than workers in smaller firms and establishments. To do this, we estimate Equation (1) using LFS data and report the coefficient β_f in Table B.3 where f represents firm size (first three columns) and establishment size (last three columns). As with the CPS data, usual hours worked are longer in larger firms and establishments. Indeed, workers in larger establishments work around 4% longer than similar workers in smaller establishments. While relative to

[1]The minimum wage in Canada is taken to be an employment-weighted average of the minimum wage across provinces as reported by Statistics Canada.

workers of firms with under 20 employees, workers in mid-sized firms work between 4 and 5% longer in larger firms.

Table B.3: Size and hours worked

	Firm Size			Establishment Size		
	(1)	(2)	(3)	(4)	(5)	(6)
20 to 99 Employees	0.064***	0.051***	0.048***	0.049***	0.038***	0.034***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
100+ Employees	0.060***	0.047***	0.041***	0.072***	0.054***	0.039***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year, Province FE	Y	Y	Y	Y	Y	Y
Demographic Controls	N	Y	Y	N	Y	Y
2-digit Industry FE	N	N	Y	N	N	Y
N	6,552,536	6,552,536	6,552,536	6,846,599	6,846,599	6,846,599
R^2	0.015	0.103	0.137	0.019	0.106	0.137

Notes: The table reports the coefficient β_f estimated from Equation (1) where the reference size category is the smallest size category. The first three columns report results where f represents firm size categories. The last three columns report results where f represents establishment size categories. All data is from the pooled LFS sample and firm size data is available starting 1998 while establishment size data is available starting 1997. Standard errors are reported in parentheses. *** indicates, respectively, statistical significance at 1% level.

Fact 2 Average wages increase with size.

Next, we estimate Equation (2) using LFS data and report the coefficient β_f, which captures the size-wage premium, in Table B.4. Consistent with the data from the US, wages in the largest firms and establishments are indeed higher than those in the smallest firms and establishments. The LFS data indicates a wage premium of around 19% for workers in firms or establishments with over 100 employees compared to those with under 20.

Fact 3 Long-hour (short-hour) penalty decreases (increases) with size.

Finally, we show that, consistent with data from the CPS, the short and long hours penalties also vary systematically by firm and establishment size in the LFS. Figure B.6 plots the relationship between hours and wages in the LFS as estimated from Equation (2). In particular, Panels (a) and (b) report the sum of the coefficients γ_h and $\theta_{f,h}$ which captures the wage penalty of working outside of the 40-44.9 hours bin by firm and establishment size, respectively. The panels show that for both firms and establishment size categories, there exist hump-shaped relationships between hours and wages. Importantly, as with the US data, the short hours wage penalty in larger firms and establishments is much more severe than the penalty in smaller firms. For example, compared to small firms, working 25 hours in large firms (with over 100 employees) results in around 15% lower wages. The analogous measure for establishment size is around 10% lower wages. Conversely, the long-hours wage penalty is much more severe in smaller firms and establishments than in larger firms and establishments. For example, compared to large firms, working 60 hours in small firms

Table B.4: The size-wage premium

	Firm Size			Establishment Size		
	(1)	(2)	(3)	(4)	(5)	(6)
20 to 99 Employees	0.097***	0.078***	0.080***	0.093***	0.077***	0.081***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
100+ Employees	0.212***	0.176***	0.185***	0.241***	0.201***	0.191***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Year, Province FE	Y	Y	Y	Y	Y	Y
Demographic Controls	N	Y	Y	N	Y	Y
2-digit Industry FE	N	N	Y	N	N	Y
N	6,552,536	6,552,536	6,552,536	6,838,698	6,838,698	6,838,698
R^2	0.146	0.273	0.342	0.159	0.284	0.348

Notes: The table reports the coefficient β_f estimated from Equation 2 where the reference size category is the smallest size category. The reference hours bin is $40 - 44.9$ hours. The first three columns report results where f represents firm size categories. The last three columns report results where f represents establishment size categories. All data is from the pooled LFS sample and firm size data is available starting 1998 while establishment size data is available starting 1997. Standard errors are reported in parentheses. *** indicates, respectively, statistical significance at 1% level.

(with under 20 employees) results in around 5% lower wages. The analogous measure for establishment size is around 3% lower wages.

(a) Firm Size (b) Establishment Size

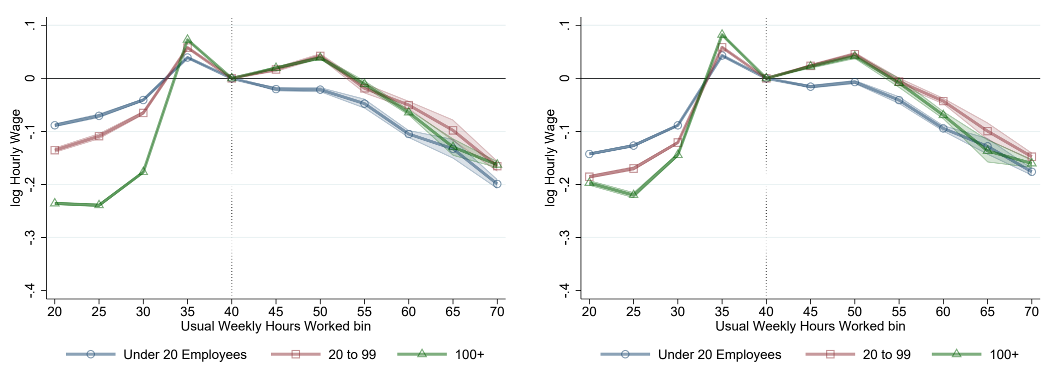

Figure B.6: The relationship between wages and hours, Canada

Note: Panels (a) and (b) report the the sum of coefficients $(\gamma_h + \theta_{f,h})$ estimated from Equation (2) using LFS data where f represents, respectively, firm size and establishment size. The reference group for usual hours worked in the regression is workers that work $40 - 44.9$ hours. The reference group for size is the smallest size category that is firms or establishments with under 20 employees. The shaded regions are the 95% confidence intervals. Data is from the pooled LFS sample.

Taken together, the analysis with the LFS data is encouraging as is confirms that our main empirical findings are not simply an artefact of the US data. Additionally, replicating our motivating facts using Canadian data is encouraging as it suggests that the US and Canadian economies are similar and may share similar fundamentals such as the substitution parameter ρ.

C Model Appendix

In this appendix, we provide additional details and discussions related to the model and quantitative analysis.

C.1 The role of taste shocks

Described in Section 3, our model features shocks to the return of workers to working for firms of differing productivity. The computational advantage of these shocks is that they help 'convexify' the occupational choice of workers by introducing additional randomness in their decision.

In particular, by assuming a Generalized Extreme Value Distribution for these shocks, the occupational choice of workers can be considered as a probability which is given by the value obtained in each occupation – net of the ϵ-shocks, relative to the aggregation of values in all other firm productivity levels,

$$H(j;v) = \frac{\exp\left(V^G(z_z,v)\right)^{\frac{1}{\sigma_\epsilon}}}{\sum_{k=1}^{J} \exp\left(V^G(z_k,v)\right)^{\frac{1}{\sigma_\epsilon}}}.$$

Having a probability as the policy function, instead of an binary indicator of 0 or 1 for choosing each occupation, smooths out the value function of workers, and help with convergence.

Existing literature has used similar "tastes" in different models of discrete choice, such as McFadden (1978) for households' location choice and Wolpin (1984) in a model of fertility. The role of taste heterogeneity in shaping wage heterogeneity between employers has previously been highlighted and modeled in Card et al. (2018). Other papers that use similar shocks in the context of occupational choice of workers are Artuç et al. (2010) and Caliendo et al. (2019).

As with Card et al. (2018), these taste shocks play a crucial role in generating a size-wage premium in our model. We interpret these taste shocks as capturing a number of data features that affect individuals' sorting into firms of different productivity and size, which are not accounted for by wages. Below we discuss several features of the labor market that motivate introducing the taste shocks in the model.

Small and large firms differ in the multi-dimensional non-pecuniary benefits they offer to workers. Studies showed that workers' heterogeneous preferences over these non-pecuniary characteristics are important in generating earning inequality (Rosen, 1986, Morchio and Moser, 2018 and Lamadon et al., 2022). Importantly, we emphasize that these non-pecuniary characteristics might differ across firm size groups. While small firms have a more friendly and less rigid work environment (Agell, 2004 and Idson, 1990), large firms might excel in some other dimensions such as safety in the work environment (Oi, 1974).

Further, there may exist logistical and technological reasons why different workers do not find it equally feasible to work in larger firms. For example, several studies have documented firms in urban areas are more productive and larger than firms in rural areas (Headd, 2000 and Melo et al., 2009). If workers are tied to a specific location, moving cost or commuting costs could contribute to the taste shocks we build into our model. There also exists cross-sector difference in firm size, with smaller firms more likely to be in the construction, services, and agriculture sector and large firms more likely to be in the manufacturing, retail, transportation, and finance sector (Headd, 2000). In this regard, the taste shock could also be interpreted as the limited transferability of sector-specific human capital.

The taste shocks we introduce are a reduced-form representation of the real-world features we discuss here. Having said this, much of the worker side's heterogeneity above can be argued to be persistent, yet our taste shocks are independently drawn each period. The reason for this is tractability and simplicity. We could generate similar implications with our model in a static model with static heterogeneity in workplace preferences.

Notice also that our taste shocks are over firms' productivity levels, even though the discussion above corresponds to the preferences or ability to work in firms of different sizes. However, firm size and productivity will be isomorphic in our model and are strongly positively correlated in the data (see, for example, Leung et al. 2008 and Bartelsman et al. 2013).

C.2 Computing welfare gains

To conduct the policy experiments in Section 5.2, we use pre- and post-policy worker values to get a simple measure of consumption equivalence (CE). In models where the value only depends on consumption, one measure of CE would be a household-specific constant H such that if consumption levels were permanently raised in the benchmark economy by a rate H, the value obtained by the household will be the same between the benchmark economy or moving to an economy with alternative policies. In our model, where hours worked and the taste shocks enter into the value flows, this measure will not have the same direct interpretation, and but will be useful to put the welfare changes in consumption terms. In particular, if we were to convert the value into an alternative derived by only consumption flows, we would have to multiply the hypothetical flows in the benchmark, with a constant $1 + H$ to reach to the value obtained in the alternative with the policy change.

$$V^k(a_0, x_0, z_{0,j}, v_0) = E_0^k\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{c_t^{1-\gamma}}{1-\gamma} - v_t\frac{l_t^{1+\phi}}{1+\phi} + \epsilon_t\right)\right] = E_0^k\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{\tilde{c}_t^{1-\gamma}}{1-\gamma}\right)\right]$$

where $k = \{\text{BM,POL}\}$ denotes the model in consideration, Benchmark or Policy, and $\{\tilde{c}_t\}_{t=0}^{\infty}$ denotes a consumption stream that would, ignoring the value of leisure and tastes, the same value as in a given model.

Then we compute the CE for a given individual state as:

$$H(a_0, x_0, z_{0,j}, v_0) = \left(\frac{V^{POL}(a_0, x_0, z_{0,j}, v_0)}{V^{BM}(a_0, x_0, z_{0,j}, v_0)}\right)^{\frac{1}{1-\gamma}} - 1.$$

In turn, the measure of winners from a policy change would be the mass of individuals for whom the coefficient H is positive.

C.3 Showing the optimality of the symmetric solution

In our analysis, we focus on a symmetric equilibrium in which all the firms in a z_j-market absorb the same share of the supply $\mu_j^s(l,x)$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$.[2] Accordingly, we need to assure that a demand schedule $\mu_{a,j}(l) = \mu_j^s(l,x)/F(z_j)$ solves the maximization problem (5) for each firm i with firm productivity z_j. To do so, we characterize firms' optimization problem and show that, under fairly relaxed constraints on labor demand, firms do not benefit from deviating from the symmetric solution.

We assume that firms are subject to a natural upper bound, $\bar{\mu}_j$, on their labor demand from any hour-efficiency pair and we parameterize this as,

$$\bar{\mu}_j = X_\mu \times \max_{l \in [\underline{l}_j, \bar{l}_j],\ x \in B_x} \mu_{a,j}(l,x)$$

where $X_\mu \in (1,\infty)$. In words, we assume that the z_j-specific upper bound is high enough so that it does not bind for any level of hours and worker efficiency in any labor market z_j in the symmetric equilibrium.

The first order conditions from firm i's problem implies that for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$,

$$xz_j\theta L_i^{\theta-1} E_i(l^\rho)^{\frac{1}{\rho}} \left[\frac{1}{\rho}\frac{l^\rho}{E_i(l^\rho)} + 1 - \frac{1}{\rho}\right] - w_j(l,x)l - \lambda_i(l,x) + \chi_i(l,x) = 0 \tag{C.1}$$

where we denote the Lagrange multipliers of the non-negativity constraint by $\chi(l,x)$, and the one for the upper bound restriction by $\lambda(l,x)$. Here, we maintain the notation that

$$E_i(l^\rho) \equiv \left(\int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} x\mu_i(l,x) l^\rho \, dl dx \right),$$

$$L_i \equiv \left[\left(\int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} x\mu_i(l,x) l^\rho \, dl dx \right)^{\frac{1}{\rho}} \left(\int_{x \in B_x} \int_{\underline{l}_j}^{\bar{l}_j} \mu_i(l,x) \, dl dx \right)^{1-\frac{1}{\rho}} \right]$$

are the labor aggregation terms that correspond to a given demand schedule μ_i. $E_{a,j}(l^\rho)$ and $L_{a,j}$ are defined analogously.

In the symmetric equilibrium, the non-negativity and the upper bound restrictions do not bind, i.e. $\lambda_{a,j}(l,x) = \chi_{a,j}(l,x) = 0$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$. Hence, the wage function in this equilibrium has to be equal to the marginal productivity of a worker from

[2] In equilibrium, all z_j-markets contain the full set of efficiency types (B_x). Without loss of generality, we proceed with the analysis by assuming that for all x-type of workers in the market z_j, there exists a common set of l, denoted by $[\underline{l}_j, \bar{l}_j]$, where there is a positive supply for worker type (l,x), i.e. $\mu_j^s(l,x) > 0$. Our analysis is still valid even if the set $\{l | \mu_j^s(l,x) > 0\}$ differs by x in the equilibrium.

each hour-efficiency pair for these firms:

$$w_j(l,x)l = xz_j\theta L_{a,j}^{\theta-1}E_{a,j}(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^\rho}{E_{a,j}(l^\rho)}+1-\frac{1}{\rho}\right],$$

for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$.

Next, we pursue whether a given firm i finds it optimal to deviate from the symmetric solution by choosing a different demand schedule. Define:

$$\begin{aligned} g^i(l,x) &\equiv xz_jL_i^{\theta-1}E_i(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^\rho}{E_i(l^\rho)}+1-\frac{1}{\rho}\right] - w_j(l,x)l \\ &= xz_j\theta L_i^{\theta-1}E_i(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^\rho}{E_i(l^\rho)}+1-\frac{1}{\rho}\right] - xz_j\theta L_{a,j}^{\theta-1}E_{a,j}(l^\rho)^{\frac{1}{\rho}}\left[\frac{1}{\rho}\frac{l^\rho}{E_{a,j}(l^\rho)}+1-\frac{1}{\rho}\right] \end{aligned}$$

Implicit in the function g^i is firm i's demand schedule μ_i. First, notice that a choice with $\mu_i(l,x) = \mu_{a,j}(l,x)$ for all $l \in [\underline{l}_j, \bar{l}_j]$ satisfies the FOCs: It implies $g^i(l,x) = 0$, with $\lambda_i(l,x)=0$ and $\chi_i(l,x) = 0$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$.

Below suppose there exists a demand schedule $\mu_i(l,x) = \tilde{\mu}_i(l,x)$ that delivers a higher level of profits than $\mu_i(l,x) = \mu_{a,j}(l,x)$. In Claim 1 and 2, we provide an analytical characterization of such a demand schedule. Guided by these results, we then examine whether there exists such an alternative demand schedule in our calibrated model where firms are better off than under the symmetric solution.

Claim 1. Suppose there exists a demand schedule $\mu_i(l,x) = \tilde{\mu}_i(l,x)$ that delivers a higher level of profits than $\mu_i(l,x) = \mu_{a,j}(l,x)$. Then, for any given $x \in B_x$, the function $g_l^i(l,x)$ associated with this demand schedule is either (i) strictly positive for all $l \in [\underline{l}_j, \bar{l}_j]$, or (ii) strictly negative for all $l \in [\underline{l}_j, \bar{l}_j]$.

Taking the derivative of g^i with respect to l, for a given x:

$$g_l^i(l,x) = xz_j\theta l^{\rho-1}\left[L_i^{\theta-1}E_i(l^\rho)^{\frac{1}{\rho}-1} - L_{a,j}^{\theta-1}E_{a,j}(l^\rho)^{\frac{1}{\rho}-1}\right].$$

The term in brackets is constant for all $l \in [\underline{l}_j, \bar{l}_j]$. Our objective is to show that for a schedule $\tilde{\mu}_i$ to achieve higher profits than $\mu_{a,j}$, the following condition has to hold:

$$A \equiv L_i^{\theta-1}E_i(l^\rho)^{\frac{1}{\rho}-1} - L_{a,j}^{\theta-1}E_{a,j}(l^\rho)^{\frac{1}{\rho}-1} \neq 0.$$

To see that, take some $\hat{x} \in B_x$. Suppose $A = 0$, which implies a constant $g^i(l,\hat{x})$ over l. If this constant value is negative, it implies that $g^i(l,x) < 0$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$ because g^i is proportional to x. That gives $\tilde{\mu}_i(l,x) = 0$ for any $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$ which implies zero output and zero profits. With the same token, if $g^i(l,\hat{x})$ over l is constant at a positive level, it implies that $g^i(l,x)$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$. $\chi_i(l,x) > 0$ and $\tilde{\mu}_i(l,x) = \bar{\mu}_j$ for any $l \in [\underline{l}_j, \bar{l}_j]$. This solution implies negative profits for any $\bar{\mu} > \max\limits_{l \in [\underline{l}_j, \bar{l}_j]}(\mu_{a,j}(l,x))$ in our calibration. The only remaining possibility to evaluate is that $g^i(l,\hat{x}) = 0$ for all $l \in [\underline{l}_j, \bar{l}_j]$, which implies $g^i(l,x) = 0$ for all $l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$. For $A = 0$ and $g^i(l,x) = 0$ to hold simultaneously, it is required that $E_i(l^\rho) = E_{a,j}(l^\rho)$ and $L_i = L_{a,j}$. This implies that $Y_i = Y_{a,j}$, that is, the output levels are the same under schedules $\tilde{\mu}_i$ and $\mu_{a,j}$. Since profits are equal to a constant fraction of output when $g^i(l,x) = 0$, the profits under the two schedules are also the same. This shows that for an alternative schedule to strictly dominate $\mu_{a,j}$, it has to imply that $A \neq 0$; hence $g^i(l,x)$ is strictly monotone over l for each $x \in B_x$.

Claim 2. Suppose there exists a demand schedule $\mu_i(l,x) = \tilde{\mu}_i(l,x)$ that delivers a higher level of profits than $\mu_i(l,x) = \mu_{a,j}(l,x)$. Then, one of the following statements must be true: Either

(i) $\exists\ l_0 \in [\underline{l}_j, \bar{l}_j]$ **such that for any** $x \in B_x$**,** $\tilde{\mu}_i(l,x) = 0\ \forall\ l < l_0$**,** $\tilde{\mu}_i(l,x) = \bar{\mu}_j\ \forall\ l > l_0$**, and** $\tilde{\mu}(l_0,x) \in [0, \bar{\mu}_j]$**, or**

(ii) $\exists\ l_0 \in [\underline{l}_j, \bar{l}_j]$ **such that for any** $x \in B_x$**,** $\tilde{\mu}_i(l,x) = 0\ \forall\ l > l_0$**,** $\tilde{\mu}_i(l,x) = \bar{\mu}_j\ \forall\ l < l_0$**, and** $\tilde{\mu}(l_0,x) \in [0, \bar{\mu}_j]$**.**

We first show that $\exists\ l_0 \in [\underline{l}_j, \bar{l}_j]$ such that $g^i(l_0,x) = 0$ for all $x \in B_x$. Take an arbitrary $\hat{x} \in B_x$. We know that $g^i(l,\hat{x})$ is strictly increasing or strictly decreasing over l. If $g^i(l,\hat{x}) < 0$ $\forall\ l \in [\underline{l}_j, \bar{l}_j]$, then the proportionality of g^i with respect to x implies that $g^i(l,x) < 0$ $\forall\ l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$. This implies $\tilde{\mu}_i(l,x) = 0\ \forall\ l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$ which gives zero output and profits. Similarly, if $g^i(l,\hat{x}) > 0\ \ \forall\ l \in [\underline{l}_j, \bar{l}_j]$, it implies that $g^i(l,\hat{x}) > 0$ $\forall\ l \in [\underline{l}_j, \bar{l}_j]$ and $x \in B_x$. This implies $\tilde{\mu}_i(l,\hat{x}) = \bar{\mu}_j$. We have shown in Claim 1 that this schedule gives lower profits than the solution $\mu_{a,j}$. Hence, we can focus on the case where $g^i(l,\hat{x})$ intercept 0 at some $l_0 \in [\underline{l}_j, \bar{l}_j]$, i.e. $g^i(l_0,\hat{x}) = 0$. From the proportionality of g^i with respect to x, this gives $g^i(l_0,x) = 0$ for all $x \in B_x$. We next prove the claim by case.

Case (i): Suppose $g^i(l,x)$ is strictly increasing over l for each $x \in B_x$. Then for all $l \in [l_0, \bar{l}_j]$, we have $g^i(l,x) > 0$, hence $\lambda_i(l,x) > 0$ according to C.1. The complementary slackness condition implies that $\tilde{\mu}_i(l,x) = \bar{\mu}_j$. Similarly, for any $l < l_0$, $g^i(l,x) < 0$ which implies $\tilde{\mu}_i(l,x) = 0$.

Case (ii): Suppose $g^i(l,x)$ is strictly decreasing over l for each $x \in B_x$. Then for all $l \in [\underline{l}_j, l_0]$, we have $g^i(l,x) > 0$, hence, $\lambda_i(l,x) > 0$, which implies $\tilde{\mu}_i(l,x) = \bar{\mu}_j$. Analogous to case (i), for any $l > l_0$, $g^i(l,x) < 0$ which implies $\tilde{\mu}_i(l,x) = 0$.

Lastly, an immediate implication from the above analysis is that for any $x \in B_x$, there exists a unique l such that $l = l_0$ and $0 < \tilde{\mu}_i(l,x) < \bar{\mu}_j$. For all other $l \in [\underline{l}_j, \bar{l}_j]$ and $l \neq l_0$, either $\tilde{\mu}_i(l,x) = \bar{\mu}_j$ or $\tilde{\mu}_i(l,x) = 0$.

Taking stock. In the light of the results above, for an alternative schedule $\tilde{\mu}_i$ to dominate the solution $\mu_{a,j}$, the only possibility is that it entails a demand equal to $\bar{\mu}$ for a compact set of hours that includes either the lowest or the highest hours in the choice set (i.e., a or b, respectively), and a demand of zero amount for the rest of the hours.

For illustration purposes, let c and d be the lower and the upper bound of the compact set that demand $\bar{\mu}$. The profits obtained by this case are:

$$\tilde{\pi}_j(c,d) = z_j \left[\left(\int_{x \in B_x} \int_c^d x \bar{\mu}_j l^\rho dl dx \right)^{\frac{1}{\rho}} \left(\int_{x \in B_x} \int_c^d x \bar{\mu}_j dl dx \right)^{1-\frac{1}{\rho}} \right]^\theta - \int_{x \in B_x} \int_c^d w_j(l,x) \bar{\mu}_j l dl dx \quad \text{(C.2)}$$

It will be useful to go further in computing Equation (C.2) to get:

$$\begin{aligned} \tilde{\pi}_{c,d} \quad = \quad & z_j \bar{\mu}_j^\theta \left(\frac{\bar{x}^2 - \underline{x}^2}{2} \right)^\theta \left\{ \left(\frac{d^{\rho+1} - c^{\rho+1}}{\rho+1} \right)^{\frac{1}{\rho}} (d-c)^{1-\frac{1}{\rho}} \right\}^\theta & \text{(C.3)} \\ & -\bar{\mu}_j \left(\frac{\bar{x}^2 - \underline{x}^2}{2} \right) z\theta L_{a,j}^{\theta-1} E_{a,j}(l^\rho)^{\frac{1}{\rho}} \left[\frac{1}{\rho} \frac{d^{\rho+1} - c^{\rho+1}}{(\rho+1) E_{a,j}(l^\rho)} + (1 - \frac{1}{\rho})(d-c) \right] & \text{(C.4)} \end{aligned}$$

For any x, we can find l_0 by solving the equation $g^i(l_0, x) = 0$, in which both L_i and $E_i(l^\rho)$ can be written as a function of l_0 and $\bar{\mu}$. If $g^i(l,x)$ is strictly increasing over l, then the profit is given by $\tilde{\pi}_j(l_0, \bar{l}_j)$. On the other hand if $g^i(l,x)$ is strictly decreasing over l, then the profit is given by $\tilde{\pi}_j(\underline{l}_j, l_0)$.[3]

Next, we evaluate the gains and losses for a firm that deviates from the symmetric solution $\mu_i(l,x) = \mu_{a,j}(l,x)$ to the alternative demand schedule $\mu_i(l,x) = \tilde{\mu}_i(l,x)$ that attains the highest profit highlighted above. In doing so, we compute the largest gain attained by any firm (of any productivity z_j) relative to the symmetric solution as:

$$\max_j \left[\frac{\max\{\tilde{\pi}_j(\underline{l}_j, l_0), \tilde{\pi}_j(l_0, \bar{l}_j)\} - \pi_{a,j}}{\pi_{a,j}} \right].$$

Notice that this formula takes the restriction parameter X_μ as given. Accordingly, we compute these profit gains and losses for each $X_\mu \in [1, 20]$. We show the results in Figure C.1. The figure shows that for relatively relaxed restrictions on the labor demand, we can sustain our symmetric equilibrium. In particular, if we limit the labor demand for each hour-efficiency pair to a maximum of 15 times the labor supply of the largest group of hour-efficiency type per firm, then the symmetric solution is the optimal solution for each firm. Accordingly, in our analysis, we assume that X_μ falls within a range $(1, 15]$ and focus on the symmetric equilibrium.

[3]In theory, where there is a continuum of l, there exists a cut-off type l_0 such that firms' demand for this type is positive but not necessarily equal to $\bar{\mu}_j$ (see Claim 2). However, in our quantitative exercise, this knife-edge condition does not hold for any hour-efficiency type in the equilibrium. That is, for any type with a positive demand in the alternative solution, their demand schedule hits the upper bound $\bar{\mu}_j$.

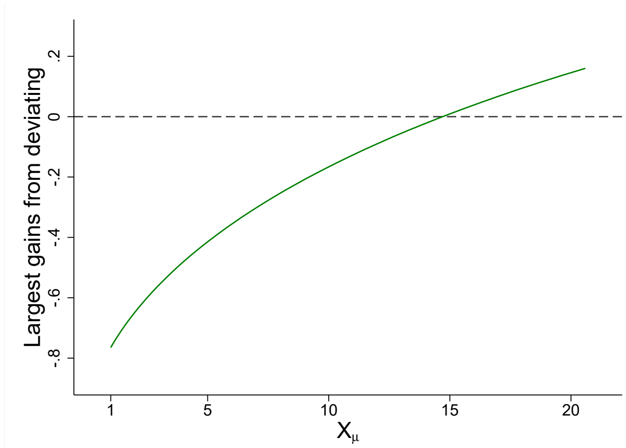

Figure C.1: Profits in the alternative solution relative to the symmetric solution

Notes: The model imposes the restriction on the demand for a given hour-efficiency pair of worker as $\mu(l,x) \leq \bar{\mu}_j$. We parameterize the upper bound as $X_\mu \times \max_{l \in [\underline{l}_j, \bar{l}_j], x \in B_x} \mu_{a,j}(l,x)$ in order to represent the tightness of the restriction with one parameter, X_μ, common across all firm productivity levels. The figure shows the rate of change in the objective function of the firms if they deviate from the symmetric solution in which all z_j-firms demand the same amount of (l,x) pair of workers.