Labor Supply and Establishment Size*

Lin Shao† Faisal Sohail‡ Emircan Yurdagul§

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

March 2021

Abstract

We find that larger establishments in Canada feature i) higher wages ii) longer worker hours and iii) smaller (larger) wage penalties for working long (short) hours. We argue that the first fact, i.e. the size-wage premium, explains the other two. We show this in a general equilibrium model of heterogeneous firms and workers. Workers willing to work longer hours self-select into larger firms because the wage premium increases earnings more with longer hours. The model generates wage penalties for short and long hours through complementarities between workers' hours. Longer hours in larger establishments, together with complementarities, leads to larger (smaller) short-hour penalties in larger (smaller) establishments, as in the data. This general equilibrium effect amplifies the increasing working hours with the firm size. Increasing hours and hourly wages with employer size contribute similarly to the earnings gap between small and large firms, both in the model and in the data.

JEL Codes: E24, J2, J31

Keywords: Labor supply, earnings inequality, production, firm size, sorting

*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 and Andres Erosa for valuable feedback. Yurdagul gratefully acknowledges the support from the Ministerio de Economia y Competitividad (Spain) (ECO2015-68615-P) and from the Ministerio de Ciencia, Innovacion y Universidades (IJC2018-038229-I).

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

‡faisal.sohail@unimelb.edu.au. Department of Economics, University of Melbourne, 111 Barry St, Carlton VIC, 3053, Australia

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

1 Introduction

Firms are typically modeled as a production function where worker hours are aggregated up and treated as an input. Firm-level variation in the distribution of working hours, such as those in the average level and dispersion, is often neglected. This paper studies the firm-level variation in the distribution of hours and wages and the interplay between the two. Our objective is to shed light on the determinants and implications of this variation empirically and theoretically.

Our study is motivated by three main facts about the distribution of hours and wages across firms.[1] The first is the well-documented *wage-size premium.* That is, the observation that average hourly wage increases with firm size (see for example Oi and Idson, 1999). Second, we show that average worker hours increase with firm size. Third, there is a wage penalty for working relatively long and short hours across all firms, and large firms exhibit a less severe penalty for working long hours and a more severe penalty for short hours. We document these facts using the Canadian Labour Force Surveys.[2]

We then develop a model with heterogeneous firms and workers that can reconcile these facts. Firms vary in their exogenous productivity and decide on static factor inputs. The production function of firms explicitly includes worker hours and allows for complementarities between workers' hours such that workers are more productive if they work a similar number of hours. Workers vary in their value of leisure and, in addition, have random preferences for working in firms of different productivities. Given these two dimensions of exogenous variation, workers choose their labor supply, savings and which firm to work for.

Despite its minimal structure, the model can generate all three motivating facts. The wage-size premium (fact 1) is generated as a result of endogenous sorting of workers based on the wages offered and their idiosyncratic tastes for firms that are ex-ante symmetric across firm productivity. The additional randomness in workers' sorting caused by varying tastes for firms makes firm employment increase less steeply with productivity and allows the marginal productivity of labor (wages) to increase with firm productivity. In the absence of

[1]Our empirical findings hold both at the firm or establishment level and our theoretical mechanisms can apply to either unit of activity. As such, throughout this paper we use the terms establishment and firm interchangeably.

[2]Analogous evidence from the Current Population Survey (CPS) in the US in the appendix section C.

a consensus explanation of the wage-size premium, we use heterogeneity in worker tastes to represent many proposed explanations in a reduced form manner. We interpret these tastes as capturing the many dimensions that affect individuals' sorting into firms of different productivity (and size) that are not accounted for by earnings. We discuss in detail the interpretations of the taste shocks in the appendix.

The positive relationship between firm size and worker hours (fact 2) is driven by the interaction of the wage-size premium and workers' sorting based on their value of leisure. The existence of the wage-size premium has two opposing effects on workers' sorting into firms of different sizes. Consider a worker with a low value of leisure, that is, a worker willing to work longer hours. This worker generates more additional income than a worker with a high value of leisure by working in a larger firm since she works longer hours. However, her utility gain from a marginal increment in income is lower than the latter type. We show quantitatively that the first effect dominates, making the lower value of leisure workers choose to work in larger firms. As a result, the average hours worked increases with firm size.

Finally, complementarities in worker hours generate a wage penalty for workers that deviate from the modal hours worked in a firm (either long or short hours). Given the longer usual hours worked in larger firms, long hours are less heavily penalized in larger firms than in smaller firms. Similarly, shorter hours are less heavily penalized in smaller firms (fact 3). This wage effect also makes the hours worked in larger firms even longer relative to smaller firms, further amplifying the positive relationship between firm size and average working hours.

Having established the model mechanisms behind our motivating facts, we turn to the importance of these mechanisms in shaping inequality in hourly wages and total earnings. Our model predicts that firm-level heterogeneity in productivity, and worker-level heterogeneity in working hours can account for the total dispersion in hourly wages. First, wages across firms reflect variation in firm-level TFP. In our model, wage dispersion shrinks by 44 percent once we control for firm TFP or size. In addition, the distribution of working hours within firms leads to wage dispersion in the presence of complementarities. This accounts for the remainder of the dispersion in hourly wages. Hence, we find that productivity differences at the firm level and working hours differences at the worker-level play a similarly important role in the observed wage dispersion. In addition to inequality in hourly wages, our model highlights a novel sorting channel through which firm TFP and worker hours jointly contribute to inequality in total earnings. In particular, we argue that higher firm productivity and wage levels in larger firms attract more of those workers with propensities to work longer hours. This amplifies the differences in total earnings between low and high productivity, and increases total earnings inequality. To be specific, our model predicts that increasing hours over firm size plays an equally important role as the increasing hourly wage over firm size in determining the gap in total earnings between small and large firms.

This paper is closely related to several strands of literature studying the interaction between firm characteristics, wages and hours worked. Among the three facts that motivate this paper, the size-wage premium has been studied extensively. Potential explanations proposed by the literature include compensation for the more unpleasant working condition, more rent sharing, the substitution of higher monitoring costs with a wage premium, and assortative matching (see Brown and Medoff, 1989 and Oi and Idson, 1999 for surveys of the literature). The way we generate the size-wage premium closely follows Card et al. (2018), where the cross-firm wage differences arise from workers' heterogeneous preference over potential employers with wage-setting powers.[3] Such heterogeneity captures data features that affect individuals' sorting into firms of different firm-productivity and size groups that are not accounted for by the current wages, such as job location or non-pecuniary benefit.[4] Like Card et al. (2018), our model generates a positive correlation between firm productivity—hence firm size—and wages in equilibrium. Nevertheless, our objective is not to highlight the exact reasons behind the size-wage premium. Instead, we generate this feature in this simple and conventional way to study its implications for the hours patterns across firm size and the within-firm variation of wages.

To our knowledge, Montgomery (1988) and Headd (2000) are the two empirical studies that touch on the relationship between the working hours patterns and firm size. Both papers 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

[3]Lamadon et al. (2019) also feature a similar setup of the random utility model.

[4]In Appendix A, we provide a detailed discussion of the empirical evidence behind the taste shocks.

increases with establishment size. Unlike these papers, we compare the full distribution of hours across establishments of different sizes, up and above the distinction of part-time and full-time workers. In line with this, our model is rich enough to capture differences between small and large firms that cannot be attributed to 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, naturally relates to the literature that documents the presence of such penalties across the economy. In particular, Yurdagul (2017) and Bick et al. (2020) show an inverse U-shaped relationship between hours worked and hourly wages, robust to controlling for relevant characteristics of workers, jobs, and sectors. This paper shows that these penalties vary with firm size, namely that larger firms have a lower penalty for long hours and a higher penalty for short hours.

Our model implications on the determinants of earnings inequality is in line with the empirical studies that show that firms have played an important role in explaining the rise in inequality over time. Song et al. (2019) study how much of the rise in earnings inequality can be attributed to the with- or between-firm variations. Their results show that an increasing variation in the average firm earnings contributed to most of the rise in earnings inequality over the past few decades. In addition, evidence from Mueller et al. (2017) suggests a positive link between an increase in the size of the largest firms and the rise in inequality in developed countries. Since our model features firm productivity heterogeneity and generates a distribution in hours and wages within and across firms, it provides a suitable framework to analyze how an increasing variation in firm productivity and the rise of superstar firms leads to changes in the within- and between-firm earnings inequality.

In the next section, we review our motivating facts. In Section 3 we outline a simple model to illustrate the role of heterogeneity in wages across firms can lead to heterogeneity in hours across firms. Section 4 details the full model. We calibrate the model in Section 5, and evaluate its quantitative predictions in Section 6. Section 7 concludes.

2 Motivating Facts

This section documents three motivating facts about the distribution of hours and wages across establishment size. First, we confirm an establishment-size wage premium. Second, we document a robust positive relationship between establishment size and average worker hours. Workers in the smallest establishments (with >20 employees) spend 3% fewer hours per week than those in the largest establishments (>500 employees). Third, we document a short-hour and long-hour penalty that are present for the universe of the establishments. However, the magnitude of the two penalties vary across establishment size categories. Notably, while long-hour penalty decreases with size, short-hour penalty increases with size.

Data Description Our analysis uses pooled data from 1997 to 2016 Canadian Labour Force Surveys (LFS), a nationally representative survey conducted monthly. We start our sample in 1997 as this is the first year that establishment size is available in the LFS.

The LFS contains detailed information on respondents' economic activity for the month in which they are interviewed. Importantly, the LFS includes information on workers' hourly earnings, their usual hours worked, and establishment size. Establishment size is reported in four categories: those with fewer than 20 employees, those with 20 to 100 employees, those with 100 to 500 employees, and more than 500 employees.[5] 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 federal minimum wage. Our final sample includes 6.23 million respondents.

2.1 Establishment Size, Hours and Wages

Fact 1 Average wages increase in establishments size.

The wage premium in large establishments and establishments has been studied extensively in the literature, see for example Oi and Idson (1999). Figure 1 plots both the unconditional

[5]The LFS also provides information on firm size and we document the relationship between firm size, hours and wages in the appendix.

and conditional average log hourly earnings by establishment size. The figure documents a size premia of around 10% to 15% between the largest and smallest size categories regardless of the usual worker hours worked.[6]

(a) Average Unconditional Wage (b) Average Conditional Wage

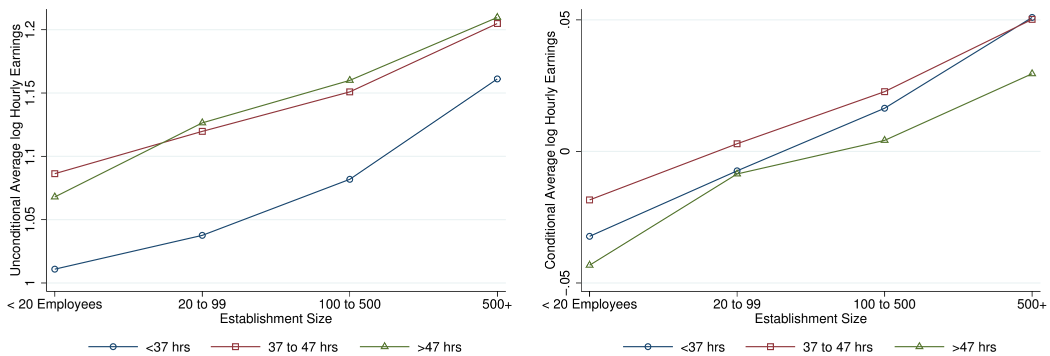

Figure 1: The Establishment-Size Premia

Note: Figure uses data from the pooled LFS sample. Panel (a) reports the unconditional average of log hourly earnings of workers by establishment size across three hour bins. Panel (b) plots the average of conditional log hourly earnings. Conditional wages are the residuals from an OLS regression of log hourly wage on 4 education bins, a quadratic in age and marital status and gender dummies as well as year, province, 2-digit industry and 2-digit occupation fixed effects.

Fact 2 Average hours increase in establishments size.

Focusing next on the relationship between size and average worker hours, Figure 2 shows the distribution of hours worked by establishment size. For clarity, we bin establishments into three categories; i) small (under 20 employees), ii) medium (between 20 and 100 employees), and iii) large establishments (over 100 employees).

Panel (a) shows the overall distribution of usual hours worked. While the median weekly hours across all establishments is between 37.5 and 42.5 hours, there are important differences in the share of short and long hours worked by establishment size.[7] This can be seen in panel (b) and (c) which reports the distribution of right and left tails of the hour distribution. Panel (b) shows that workers in small establishments are much more likely to work shorter (<37.5)

[6]Table D.6 in the appendix shows that this result is robust to controlling for observable individual characteristics such as industry and occupation fixed effects. With these controls, workers in establishments with over 500 employees earn around 10% more than those in establishments with fewer than 20 employees.

[7]The LFS reports hours worked to one decimal place. The majority of workers (50.5%) in the pooled LFS report working exactly 40 hours, a significant share of workers (8%) report working exactly 37.5 hours, which is why we group these hours in a single category.

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

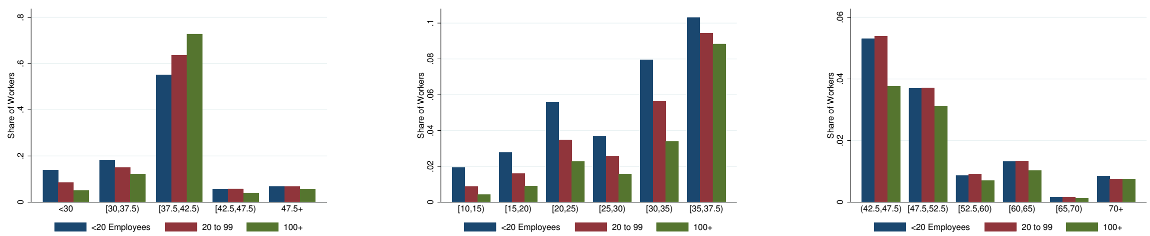

Figure 2: Distribution of Working Hours by Establishment Size

Note: Figure uses data from the pooled LFS sample.

hours than their counterparts in larger establishments, while only around 11% of employees in large establishments work 30-37.5 hours, the analogous share in small establishments is around 18%. Panel (c) shows that employees in small and medium-sized establishments are more likely to work between 42.5 and 47.5 hours relative to their counterparts in large establishments. Very few employees work longer than 47.5 hours, and for those that do, there is very little difference in the distribution by establishment size. This is not surprising as, under normal circumstances, the maximum time an employee may work each week is 48 hours.[8]

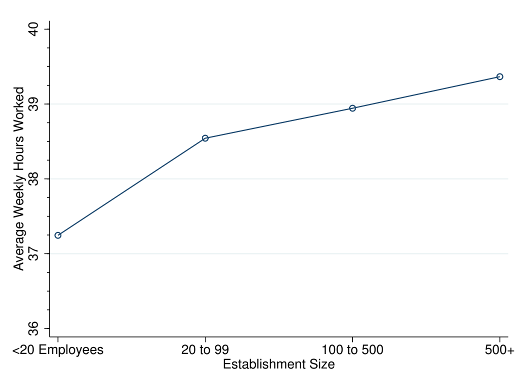

Figure 3: Average Hours by Establishment Size

Note: Figure uses data from the pooled LFS sample.

As suggested by Figure 2, there is a positive relationship between average hours worked and establishment size. This can be seen in Figure 3 which shows that, on average, workers in the largest establishments work 1 hour longer than workers in the smallest establishments. While suggestive, this cross-sectional average does not control for confounding factors that

[8]To exceed this maximum, the Minister of Labour must make an exception.

might impact both establishment size and hours worked. For example, certain industries might have both a larger optimal scale and require long hours. Table 1 reports the results from a regression on hours worked and several important worker-level controls, including a quadratic term in age, four education bins, gender, marital status, as well as controls for the province, year, industry, and occupation fixed effects. When controlling only for worker characteristics, the table shows that relative to establishments with fewer than 20 employees, workers in establishments with 20 to 99, 100 to 500, and more than 500 employees work around 3%, 5%, and 6% longer hours, respectively. Controlling for industry and occupation explains much of the difference in hours worked among large establishments but still implies that workers in establishments with over 20 employees worker around 3% longer than workers in smaller establishments.[9]

Table 1: Elasticity of Establishment Size and Hours Worked

Establishment Size (Employees)	(1)	(2)	(3)
20 to 99	0.034***	0.032***	0.031***
	(0.000)	(0.000)	(0.000)
100 to 500	0.048***	0.033***	0.034***
	(0.000)	(0.000)	(0.000)
More than 500	0.058***	0.031***	0.031***
	(0.000)	(0.000)	(0.000)
Demographic Controls	Y	Y	Y
Year, Province FE	Y	Y	Y
2-digit Industry FE	N	Y	Y
2-digit Occupation FE	N	N	Y
N	6,228,489	6,228,489	5,917,665
R^2	0.125	0.159	0.187

Notes: The table reports the estimated coefficient β_e from the following OLS regression:

$$\log(h_i) = \alpha + \sum_e \beta_e \mathbb{I}_{ei} + \gamma X_i + \epsilon_i$$

where h_i is the usual hours worked by an individual $i.X_i$ is a vector of individual level controls which include a quadratic in age, 4 education bins, gender, marital status as well as controls for province, year, industry and occupation fixed effects. The variable $\mathbb{I}_{ei}$ is an indicator which is equal to 1 if i is employed in an establishment of size e. The reference establishment size category is establishments with under 20 employees.

[9]Table D.7 in the appendix show that the data pattern still holds and is somewhat stronger without controlling for gender and marital status.

<u>Fact 3</u> Long-hour penalty decreases, and short-hour penalty increases in establishment size.

Next, we study the cross-sectional relationship between wages and hours worked by establishment size. Figure 4 plots the average wages of workers by hours worked and establishment size. A number of salient features are present in the figure. First, and consistent with fact 1, the wage-size premium is evident all over the working hour distribution. Second, there is a peak in average log hourly wages for all establishment sizes between at least 37.5 hours and less than 40 hours. Third, and consistent with the literature, we find a hump-shaped relationship between hours and earnings across establishment sizes. The figure also suggests that the decline in average wages as hours worked increase is steeper in small relative to medium and large establishments. However, these simple averages do not control for observable characteristics of workers or the establishment such as industry.

(a) Average Unconditional Wage (b) Average Conditional Wage

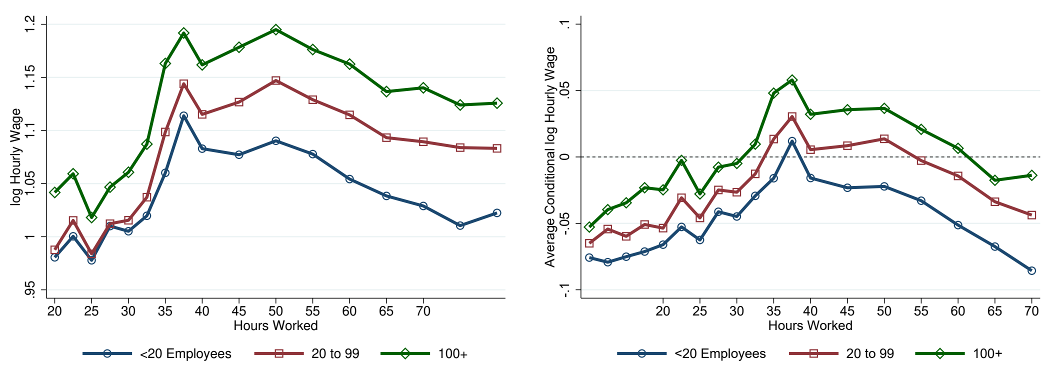

Figure 4: Average Wages by Hours Worked and Establishment Size

Next, we study the relationship between wages and hours worked while controlling for these observable characteristics. To do so, we estimate the following regression:

$$\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i,$$

where w_i is the log hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works h hours with the reference group being those that work at least 37.5 hours and less than 40 hours (i.e., the peak of the wage-hours relationship in Figure 4. Control variables captured in X_i include a quadratic term in age, four education bins, race, marital status, and controls for the province, year, 2-digit industry fixed effects.

Figure 5 plots the coefficient β_h for different establishment sizes. First, we confirm the hump-shaped pattern in the wage-hours profile previously documented in Yurdagul (2017) and Bick et al. (2020). Relative to working around 39 hours, there is a penalty for working either shorter and longer hours. We also find that the long-hour penalty is higher for smaller establishments by dis-aggregating across establishment sizes. Indeed, working over 70 hours in a small establishment leads to a 10% penalty relative to a 7.5% penalty in medium or large-sized establishments. Similarly, there is a 4% penalty for working 45 to 49 hours in small establishments compared to around 2% penalty in larger establishments. That is the long-hour penalty declines with establishment size. There also appears to be some evidence of a less severe short-hour penalty in small establishments, particularly for workers working less than 30 hours.[10]

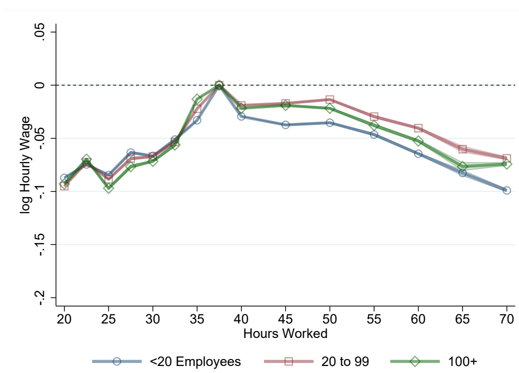

Figure 5: The Relationship between Wages and Hours by Establishment Size

Note: Figure uses data from the pooled LFS sample and plots the coefficient from a regression of log hourly earnings on hours worked dummies and demographic controls. The regression equation is $\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i$ where w_i is the hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works h hours with the reference group being those that work at least 37.5 hours and less than 40 hours. The shaded regions are the 95% confidence intervals.

Table 2 also illustrates the long hours penalty by regressing the difference in conditional wages from the peak average conditional wage at 37.5 hours on the difference in hours worked by workers from 37.5 hours. In particular, the table reports the coefficient β from the following regression separately for when workers work short hours (<37.5) and long hours (>37.5).

[10]This pattern is robust to different sets of controls (see Figure D.5 and Figure D.6 in the appendix).

$$\log(\tilde{w}_i) - \log(\tilde{w}_{h^*}) = \alpha + \beta(h_i - h^*) + \gamma X_i + \epsilon_i,$$

where $\log(\tilde{w}_i)$ is the conditional wage of individual i, h_i is hours worked, h^* is 37.5 and X_i a vector of control variables as defined above.

Consistent with Figure 5, the results from this regression show that working an additional hour when working below 37.5 hours leads to an increase in (conditional) wages in establishments of all sizes with a slightly larger increase in medium-sized establishments. Working an additional hour when already working long hours leads to a decrease in (conditional) wages, and this decrease is largest in small establishments.

Table 2: Change in Wage Penalty when Working Long and Short Hours

	Establishment Size			
Change in Wage Penalty	$<$20 Employees	20 to 99 Employees	100-499 Employees	500+ Employees
Short Hours	0.0024***	0.0028***	0.0027***	0.0017***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Long Hours	-0.0018***	-0.0011***	-0.0013***	-0.0011***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Before presenting our full model that reconciles the three facts highlighted in this section, we provide a simple example to highlight how the well-known size-wage premium, Fact 1 above, can lead to increasing hours with firm size.

3 A Simple Model

Consider a static partial equilibrium model with two firms paying constant hourly wage w_1 and $w_2 = w_1(1+\mu)$, where μ is a constant positive parameter. Suppose there is a continuum of workers that are only different in their fixed labor supply $l > 0$. Their consumption is

$$c_j(l) = w_j l + X$$

where $j \in \{1, 2\}$ is the index of the employer firm, and $X > 0$ is an endowment constant across workers. Workers' preferences over consumption are represented by $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$.

In addition to the utility from consumption, workers also have a taste for working in firm 1, represented by ϵ, which has a cumulative distribution function $F(\epsilon)$. Accordingly, a worker chooses to work in firm 2 if

$$u(w_1 l + X) + \epsilon \leq u(w_2 l + X)$$

is positive. Then the probability that a worker chooses firm 2 over firm 1 is

$$F\left[N(l; X, \gamma)\right]$$

where $N(l; X, \gamma) \equiv u(w_2 l + X) - u(w_1 l + X)$. For μ small enough, we have

$$N(l; X, \gamma) \approx u'(w_1 l + X)\mu w_1 l$$

If this term increases with l, then the higher-wage firm would have longer hours on average. The derivative of N with respect to hours is $N'(l; X, \gamma) =$

$$u'(w_1 l + X)\mu w_1 \left[\frac{u''(w_1 l + X) w_1 l}{u'(w_1 l + X)} + 1\right] = u'(w_1 l + X)\mu w_1 \left[-\gamma + 1 + \gamma \frac{X}{w_1 l + X}\right]$$

which is positive if and only if

$$-\gamma\left(1 - \frac{X}{w_1 l + X}\right) + 1 > 0. \tag{1}$$

In words, the probability of choosing the higher-wage firm increases with working hours if and only if the constant endowment is large enough or the constant relative risk aversion is low enough.

The intuition of the above result is simple. There are two channels through which hours affect workers' willingness to choose the high-wage firm. First, the longer is the hours; the more additional income is generated by working in the high-wage firm. This channel makes the probability of choosing the high-wage firm increase with the hours of a worker. Second, the longer is the hours, the marginal utility obtained by a given increase in income is lower. This channel makes the probability of choosing the high-wage firm decrease with the hours of a worker. If the consumption levels are higher (i.e., higher X), or the risk aversion parameter is lower (i.e., lower γ), the first channel dominates, leading to longer hours in the high-wage firm relative to the low-wage firm.

This simple model provides the intuition for the sorting of workers with different desired hours into firms with different wages. In the next section, we present our full model, in which larger firms pay higher wages in equilibrium. Consistent with the intuition of our simple model, this makes average hours longer in larger firms.

4 Model

The agents in the economy are firms and households. We go over their decisions separately.

4.1 Firms

There is a continuum of firms with unit mass. Production of all the firms in the economy can be represented by $Y = z(K^{\alpha}L^{1-\alpha})^{\eta}$, where K denotes capital, L denotes the effective labor, z is the firm's TFP, and η gives the span-of-control. Firms are different in their exogenous TFP, a discrete random variable distributed with $\lambda(z)$, and can take J different values. In what follows, we will denote the index of a firm productivity level by j.

Effective labor of a firm depends on the working hours distribution of the workers employed in a firm: L denotes the aggregate hours of work in the firm. As in Yurdagul (2017), we allow for complementarities between hours of workers:

$$L = \left(\int_{i \in N} l_i^{\rho} di \right)^{\frac{1}{\rho}} \left(\int_{i \in N} 1 di \right)^{1-\frac{1}{\rho}}, \tag{2}$$

where N is the set of workers, and $\{l_i\}_{i \in N}$ is their hours worked, with $l_i \in [0, 1]$, $\forall i \in N$. In order to abstract from indices of workers, one can rewrite the aggregation in terms of measure of workers employed with each level of hours worked:

$$L = \left(\int_0^1 \mu(l) l^\rho dl \right)^{\frac{1}{\rho}} \left(\int_0^1 \mu(l) dl \right)^{1-\frac{1}{\rho}}, \tag{3}$$

where $\mu(l)$ is the measure of workers working l hours.

Firms cannot distinguish between workers with a given labor supply, and there is a market price for each productivity level of firms hiring a worker with a given working hours, $w_j(l)$. They also face competitive market prices for renting capital at rate $r+\delta$. They maximize static profits choosing how much capital to rent, K; the measure of workers to work l hours:

$$\pi_j = \max_{K,\mu} \quad Y - (r+\delta)K - \int_0^1 w_j(l)\mu(l) l dl \tag{4}$$

$$s.t. \quad Y = z\left[K^\alpha \left[\left(\int_0^1 \mu(l) l^\rho dl \right)^{\frac{1}{\rho}} \left(\int_0^1 \mu(l) dl \right)^{1-\frac{1}{\rho}} \right]^{1-\alpha}\right]^\eta . \tag{5}$$

The optimal capital and the measure of labor is denoted by K_j and $\mu_j(l)$. Each firm is owned by all of the households in the economy with infinitesimal shares for each household.

Hump-shaped wage schedule. The first order conditions of a firm with productivity z imply that the equilibrium wages satisfy:[11]

$$w_j(l) \quad = \quad (1-\alpha)\eta \left(\frac{\alpha\eta}{r+\delta} \right)^{\frac{\alpha\eta}{1-\alpha\eta}} \left(\frac{z^{\frac{1}{1-\eta}} \lambda(z)}{L_j^s} \right)^{\frac{1-\eta}{1-\alpha\eta}} E_j(l^\rho)^{\frac{1}{\rho}} \left[\frac{1}{\rho} \frac{l^{\rho-1}}{E_j(l^\rho)} + (1-\frac{1}{\rho}) l^{-1} \right], \tag{6}$$

where on the right hand side we use aggregate statistics from all the workers working for a firm productivity j. In particular, $E_j(l^\rho)$ is the expected l^ρ for these workers, $L_j^s = E_j(l^\rho)^{\frac{1}{\rho}} M_j^s$

[11] The first order conditions of firms' problem imply:

$$\alpha\eta \frac{Y}{K} = r+\delta$$

$$(1-\alpha)\eta \frac{Y}{L} E_j(l^\rho)^{1/\rho} \left[\frac{1}{\rho} \frac{l^\rho}{E_j(l^\rho)} + 1 - \frac{1}{\rho} \right] = w_j(l) l, \text{ if } \mu_j(l) > 0$$

is the labor aggregation among these workers, where M_j^s is the mass of workers with that wage schedule.

Equation (6) shows that a worker's wage depends on her hours of work relative to fellow workers with the same skill level. The maximum hourly wage is achieved at: $l_j^* = E_j(l^\rho)^{\frac{1}{\rho}}$. Wages decrease as working hours get further away from l_j^*.

4.2 Workers

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

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{c_{it}^{1-\gamma}}{1-\gamma} - v_{it} \frac{l_{it}^{1+\phi}}{1+\phi} \right]. \tag{7}$$

For all workers, the value of leisure follows a Markov process $\Gamma_v(v'|v)$.

In addition to the opportunity to decide on the occupation and the value of leisure, a worker receives shocks to the value received 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 productivities. This vector follows a distribution $F(\boldsymbol{\epsilon})$ and is drawn independently every period. These shocks capture the workers' taste for different occupations and other factors locating a worker into a firm not featured in our model. Introducing such taste shocks are useful for two aspects. First, they prevent workers' sorting into firm-productivity groups from completely equalizing wages across these. By making firm size increase less steeply with firm productivity, they leave room for more productive firms to pay higher wages, while also employing more workers, which delivers the size-wage premium. Second, they help convexify workers' problem by turning their policy functions for choosing different firm-productivity levels into ex-ante probabilities. This facilitates convergence of our model, which we explain in our appendix.[12]

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

[12] The literature has long used similar "tastes" in different models of discrete choice, such as McFadden (1978) in households' location choice and Wolpin (1984) in fertility. The role of taste heterogeneity in shaping the wage heterogeneity between employers has previously been highlighted and modelled 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). It is also common in recent studies to use similar shocks to overcome computational challenges in different models of dynamic discrete choices, such as Dvorkin et al. (Forthcoming) in the context of sovereign default.

she will work for, she chooses the labor supply and the savings.

Accordingly, her value function is:

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

where the value conditional on matching with a firm of productivity j is:

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

The FOC for the static labor supply problem we get:

$$\left\{A_j \left[\frac{1}{\rho}\frac{l^\rho}{E_j(l^\rho)} + (1 - \frac{1}{\rho})\right] + \Pi\right\}^{-\gamma} \frac{A_j}{E_j(l^\rho)} = vl^{1+\phi-\rho} \tag{8}$$

where

$$A_j \equiv (1-\alpha)\eta \left(\frac{\alpha\eta}{r+\delta}\right)^{\frac{\alpha\eta}{1-\alpha\eta}} \left(\frac{z_j^{\frac{1}{1-\eta}}\lambda(z_j)}{L_j^s}\right)^{\frac{1-\eta}{1-\alpha\eta}} E_j(l^\rho)^{\frac{1}{\rho}}$$

We denote the policy function of workers for the occupational choice by $\mathbf{o}(a, v)$, and the ones for the asset and labor supply choices by $\mathbf{a}(a, j, v)$ and $\mathbf{l}(a, j, v)$.

We assume that the interest rate r is exogenous.

Stationary general equilibrium. A stationary general equilibrium is a set policy functions for firms, K_j and $\mu_j(l)$; policy functions for workers $\mathbf{a}(a, z, v)$, $\mathbf{l}(a, z, v)$ for $j \in \{1, .., J\}$ and $\mathbf{o}(a, v)$; prices r, $w_j(l)$, and a time-invariant distribution of workers $\varphi(a, z_j, v)$ over wealth (a), type of employer (z_j), and value of leisure (v) such that:

(i) Policy functions solve the problem of consumers.

(ii) K and μ solve the problem of firms.

(iii) Asset markets clear. The total capital used by the firms is equal to the total wealth in the economy.

(iv) Labor markets clear. Total measure of workers demanded by all firms for each level of productivity z_j and working hours $l \in [0, 1]$ is equal to the corresponding labor supply.

(v) Profits received by households are equal to the sum of profits of each firm.

5 Calibration

Functional forms. We assume Pareto distribution for firm productivity, with a shape parameter λ, and normalize the threshold productivity at $\underline{z} = 1$. We approximate this distribution with a productivity grid of five points.[13]

We assume that the ϵ-shocks affecting workers' value in each firm follow a Generalized Extreme Value distribution:[14]

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

The distribution gives us two parameters: ρ, dictating the correlation between the components of the $\boldsymbol{\epsilon}$ vector; and σ, determining the variance of these shocks. We assume ρ_ϵ equal to 1 for simplicity.

Even though we assume discrete Markov process for the value of leisure shocks, we set the parameters for this process so that it resembles features of an AR(1) process as:

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

This way, we boil down the corresponding parameters to ρ_v, σ_v and v_0. We use the Tauchen algorithm to connect this hypothetical AR(1) process to the Markov process we assume for the model.

[13]We set the grid points such that the theoretical probability of drawing a point between two grids reduces to a half when the grid indices increase by one.

[14]In the appendix we provide the implication of the ϵ-shocks and the particular distribution we assume for the policy functions of workers.

Parameters calibrated outside the model. We calibrate our model targeting features of Canadian firms and workers in the year 2015. We take the model period as one year. To the extent possible, we set the values for the parameters outside the model, following standard values in the literature (see the first panel of Table 3). In particular, we assume a risk aversion parameter (γ) at 1.5 and the inverse of the Frisch-elasticity of labor supply (ϕ) at 2. Among the standard production parameters, we set the capital share (α) at 0.33, depreciation rate (δ) at 0.06, and the span-of-control parameter (δ) at 0.85.

Parameters targeting features in the data. The remaining parameters are calibrated to match specific targets in the data. We assume an exogenous number of firms and workers. Hence the average size of firms in our model is directly determined by the ratio of former to latter. Accordingly, we will not study the level of firm size but will focus on the variation in firm size. We calibrate the shape parameter of firms' productivity distribution (λ) to match the standard deviation in log-employment.

We calibrate the persistence of the value of leisure shocks (ρ_v) to match the autocorrelation in log-hours of workers and the volatility (σ_v) to match the standard deviation in log-hours. The level parameter (v_0) is set to match the average hours of workers.

The preference shocks for working in different productivity levels of firms generate noise in workers' occupational choice and prevent the sorting from being entirely about the pecuniary returns. Hence, a larger size of these shocks makes the wage profiles steeper across firm productivity and size groups. Accordingly, we set σ_ϵ to match the wage ratio of the largest to smallest firms in the data.

Finally, we set the substitutability between hours of each worker (ρ) to 0.1 for the main analysis. In the following section, we will study the implications of this parameter for the rest of the moments. The model calibration is summarized in Table 3 and 4.

Table 3: Parameters

Parameter	Value	Basis
	Panel A: Outside the model	
γ	1.5	Risk aversion
ϕ	2	Inverse of labor supply elasticity
α	0.33	Capital share
δ	0.06	Depreciation rate of capital
η	0.85	Span-of-control
$\underline{z}$	1	Normalization
	Panel B: Calibrated	
λ	2.8	Std. log employment
v_0	4.2	Average weekly hours
σ_v	0.55	Std. log hours
ρ_v	0.7	Autocorr. log hours
ρ	0.1	Avg. log wage at peak - at 20 hrs
σ_ϵ	0.1	Size-wage gap

Notes: Panel A gives the parameters that are set following the standard values in the literature. Panel B includes the parameters that are calibrated to match specific data features with the model. The average and standard deviation of usual weekly hours worked is from the LFS sample. The difference between log wage at 45 and 20 hours is computed as the coefficient from a regression of log hours on hours worked and observables as described in Section 2.

Table 4: Targeted moments

Moment	Data	Model
Std. log employment	1.20	1.25
Average weekly hours	38.1	38.5
Std. log hours	0.243	0.247
Avg. log wage at peak - 20 hrs	0.1	0.1
Wage premium, fifth vs. first size quintile	20%	20%

Notes: The data moments are calculated using the 2015 LFS data.

6 Quantitative predictions of the model

This section presents the model's implications and discusses how they fare with the observed patterns. We start by showing how the model performs in terms of our motivating facts highlighted in Section 2, and then discuss the features of the model relevant for this comparison.

Reconciling the motivating facts. In our model, firms' only state variable is their productivity; hence firm size maps one-to-one to firm productivity. Since we use a firm-productivity grid of five points in our computations, we have five different size levels observed in our simulations.[15] Figure 6 shows the average wage and average hours against the firm size in our model. The dashed line highlights that the model implies that larger firms pay higher wages, as they do in the data. Accordingly, the model replicates our first motivating fact.

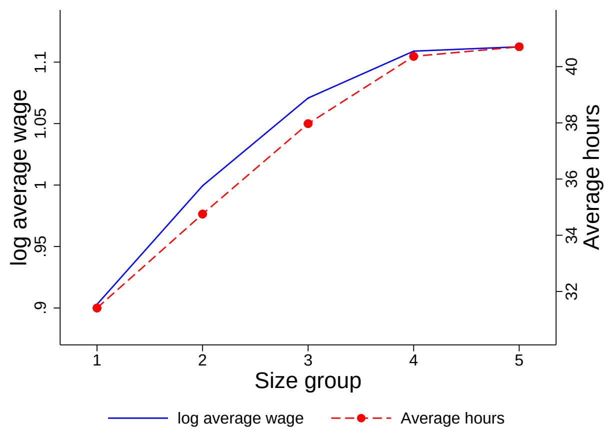

Figure 6: Wages and hours over size in the model

Note: The solid line gives the logarithm of average wage for each size group in the model. The dashed line gives the average weekly working hours in each size group. Each size group corresponds to one firm TFP grid point used in the model solution.

The solid line in Figure 6 shows that the model can also reconcile the increasing average hours with firm size observed in the data. The increase in average hours from the smallest to largest firms is as high as 8 hours per week, larger than the 3 hours observed in the data.

Figure 7, in turn, shows that the model generates our third, and final, motivating fact. In particular, for all size groups of firms, the model features a hump-shaped relationship between hours and wages. The maximum wages are observed close to the average hours for each size category, and there is an increasing penalty as a worker's hours deviate from the usual in the firm. Since the average hours increase with firm size, the penalty for working long hours decreases with firm size, and that for working short hours increases with firm size.

[15] In the data the four size categories are $1-20$, $21-99$, $100-500$, and 500+. If we normalize the employment in the lowest TFP grid to 10 (middle point of the smallest employment category in the data), the employment in the rest of the TFP levels of firms are 43, 170, 564 and 1450. Accordingly, the first four TFP grids have employment levels falling into the corresponding employment categories in the data. Nevertheless, this mapping will be more directly executed in future versions of our paper.

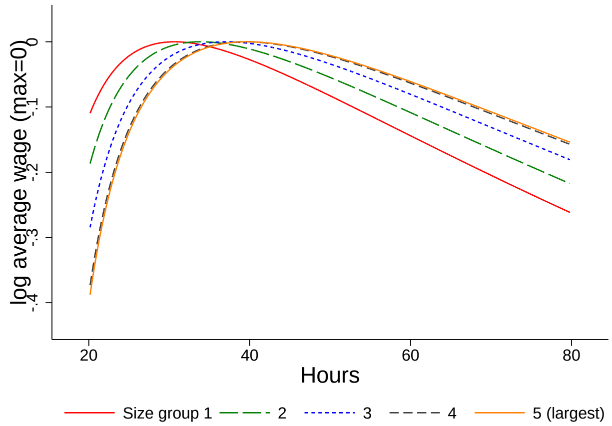

Figure 7: Wage-hours relationship and firm size

Note: The figure gives the logarithm of the wage function $w_z(l)$, in difference to its maximum level, against working hours. Each line corresponds to a different size group of firms, which in turn, corresponds to one firm TFP grid point used in the model solution.

How does the model account for the motivating facts? Through the lenses of our model, the first motivating fact, i.e., the size-wage premium, lies behind the other two facts. Accordingly, we start with discussing the model mechanism that generates the increasing wages with firm size.

In the absence of any preferences for firms of different productivities, workers would flow into higher paying firms until wages are equalized across firms. The key feature that prevents this from happening is workers' random preferences for firms (i.e., the taste shocks) as they introduce noise to workers' sorting. Larger taste shocks make wage differences less determinant in workers' sorting decisions; therefore, the equilibrium can sustain a larger wage gap between high and low productivity firms. We highlight this mechanism in Figure 8 panel (a), where we compare the size-wage patterns in our benchmark parameterization with those in the two alternatives with smaller and larger taste shocks. In particular, the variance of these shocks, σ_ϵ, is set at 0.01 and 0.2 in these two alternatives, which correspond to 10 percent and 200 percent of the benchmark value, respectively. The figure shows how the wage patterns become flatter as we reduce the size of these shocks.[16]

Since the primitive that leads to the size-hours patterns in our model is the size-wage premium, features that affect the latter pattern will affect the size-hours patterns. As men-

[16]These shocks are also useful in achieving convergence in our model solution, as they convexify a discrete choice problem. Hence we cannot show the features of the model when shut down these shocks completely.

(a) Size-wage patterns (b) Size-hours patterns

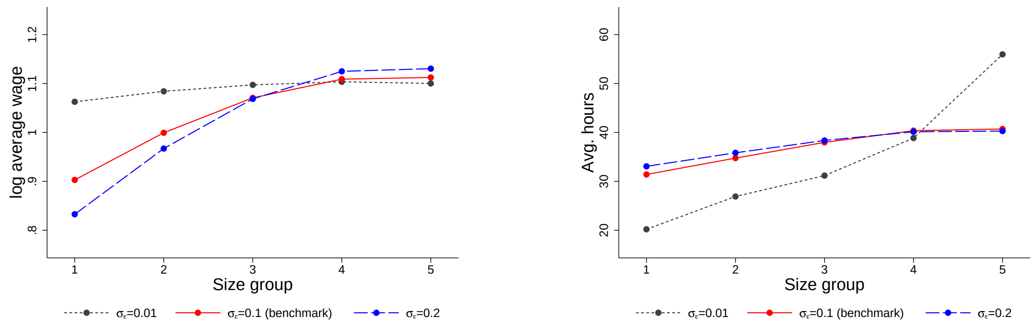

Figure 8: Size-wage and size-hours patterns, role of random preferences

Note: The left panel gives the logarithm of average wage for each size group in the model. The right panel gives the average weekly working hours in each size group. Each size group corresponds to one firm TFP grid point used in the model solution. For each figure, we show the results for the benchmark ($\sigma_\epsilon = 0.1$), together with the alternatives setting the volatility of the taste shocks equal to 10 percent of that in the benchmark ($\sigma_\epsilon = 0.01$), and to double of the value in the benchmark ($\sigma_\epsilon = 0.2$). In the alternative computations, we do not recalibrate any other model parameters.

tioned earlier, the random preferences for different productivity firms in our model are the reason behind the size-wage gap. Accordingly, there is a channel through which the smaller size for these shocks, as they decrease the size-wage premium, will flatten the increasing patterns of average hours with size. On the other hand, with smaller taste shocks, pecuniary motives become more determinant in workers' sorting across jobs, which makes the hours increase more steeply with size, for a given size-wage premium. In Figure 8 panel (b), we see that the second effect dominates, with the hours increasing more steeply with firm size when the taste shocks are smaller.

The model generates the hump-shaped wage profiles across working hours through complementarities between workers' hours. This feature maximizes the marginal productivity of an individual worker's hours when she works the same hours as the rest of her production unit. The hours above this level are not accompanied by co-workers; hence, the worker's marginal productivity diminishes for the extra hours. On the other hand, 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 7 illustrates this clearly, by plotting the wage schedule faced by a worker in a firm of median productivity. As working hours become less complementary (higher ρ), the link between hours within a unit becomes weaker, and the short and long hours penalties become smaller.

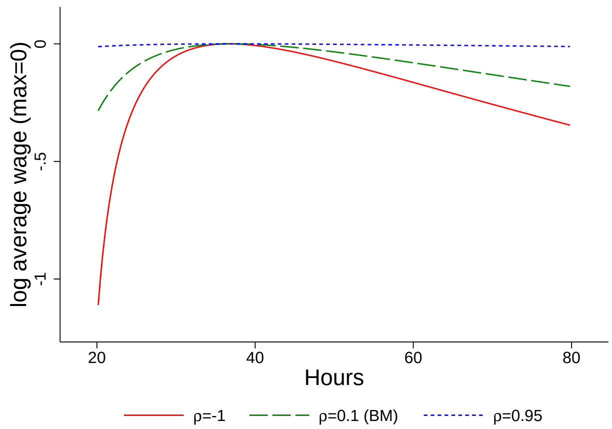

Figure 9: Wage-hours relationship and firm size; role of complementarities

Note: The figure gives the logarithm of the wage function $w_z(l)$, in difference to its maximum level, against working hours. We plot this function for the median firm TFP. The three lines correspond to the benchmark model ($\rho = 0.1$), and the alternatives with high and low complementarity between workers' hours ($\rho = -1$ and $\rho = 0.95$, respectively). In the alternative computations, we do not recalibrate any other model parameters.

Accounting for income inequality. Here, we discuss the implications of our model for the earnings inequality. First, we focus on the earnings gap between workers of small and large employers. In our model, a worker's income increases with the size of her employer for two reasons. The first is the size-wage premium, meaning that larger firms pay higher hourly wages. This would imply that the labor earnings would increase with size even if the average hours worked were the same between the firms. The second reason is that average hours also increase with firm size. The sorting of long-hour workers into larger firms contribute to the positive slope of total earnings across firm size. In figure 10 we study how these two channels contribute to the overall steepness of this size-earnings pattern. The solid blue line in the figure shows the increase in average log earnings with firm size in our model relative to earnings in the smallest size group. We see that between the smallest and largest size category, earnings increase by 47 percent. In order to understand the contribution of each of the two aforementioned channels in driving this inequality, we conduct two exercises. First, we compute a hypothetical "fixed hours" measure of earnings by taking each worker's actual wage and multiplying it by the average hours in the lowest size category. This measure is shown by the red short-dashed line. Imposing fixed hours across workers leads to a much less pronounced increase in earnings of only 21 percent. Second, we construct an analogous "fixed wage" measure of earnings which multiplies the actual hours of workers with the average hourly wage in the lowest size category. This measure is shown by the green long-dashed line. As with fixed hours, fixing wages leads to a smaller increase in earnings across firm size categories of only 26 percent. Taken together, these results show that the two channels – hours and wages – contribute similarly to the earnings gap between large and small firms accounting for 21 and 26 percent, respectively, of the 47 percent increase in total earnings.

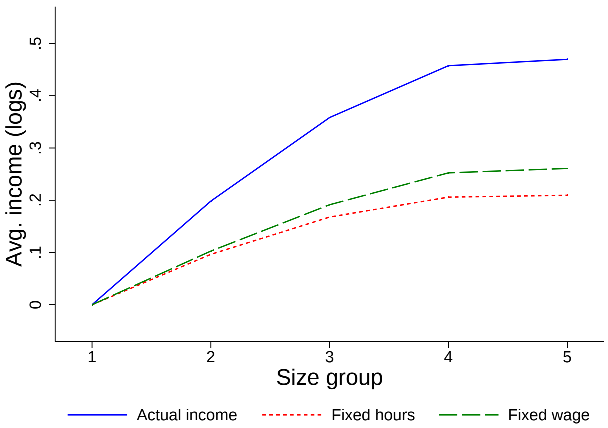

Figure 10: Earnings inequality across firm size groups

Note: The solid blue line gives the logarithm of average wage for each size group in the model, in difference to the smallest size group. The red short-dashed line gives the same for the counterfactual in which the hours are fixed at the mean of the smallest size group. The blue long-dashed line gives the same for the counterfactual in which the hourly wages are fixed at the mean of the smallest size group.

Next, we turn our attention to the earnings dispersion in general, beyond the earnings gap between small and large employers. Through the lens of our model both the firm productivity, and the level of hours worked have explanatory power in the wages of worker. Workers' wages increase with the productivity of their firm, because workers' sorting into firms is not entirely about the wages they earn. This prevents wages from equalizing between firms of different productivity, leaving a margin for a positive relationship between hourly wages and firm productivity. In addition, each firm features a distribution of wages due to the variation of hours worked. In particular, in each firm, workers earn higher wages the closer they are to the usual hours worked in the firm.

In Table 5, we highlight the role of firm productivity and hours worked in explaining the wage dispersion in our model. We do this by computing the within-group standard deviation in log-wages separately for each firm size group, each working hours bin, and each firm size-working hours bin. Since firm size is isomorphic to firm productivity in our model, each firm

Table 5: Accounting for earnings dispersion

Economy	Std. log-wages			
	Overall	Within size group	Within hours bin	Within size-hours bin
Benchmark	0.075	0.042	0.052	0.004
Low complementarity	0.066	0.005	0.061	0.001
	Std. log-earnings			
	Overall	Within size group	Within hours bin	Within size-hours bin
Benchmark	0.285	0.248	0.057	0.017
Low complementarity	0.356	0.333	0.069	0.023

Notes: The upper panel presents the standard deviation of hourly wages (in logs) for different groupings in the model. The lower panel gives the same for total earnings (hours multiplied by hourly wages). The second column of the table gives the standard deviations in a given model. The third column computes this statistic for each size (or TFP) level of firm separately, and takes the average in the economy weighting each size level by the number of workers employed in that group. The fourth column repeats the same, but grouping workers by their weekly working hours bin, and the fifth column groups workers by their firm size-working hours bin. Weekly hours bins are $\{20-, 20-22, ..., 72-74, 74+\}$. We show these statistics for the benchmark model ($\rho = 0.1$), and the alternative with low complementarity between workers' hours ($\rho = 0.95$). In this alternative, we do not recalibrate any other model parameter.

size group corresponds to a firm TFP grid. Weekly hours bins are $\{20-, 20-22, ..., 72-74, 74+\}$. Once computing this within-dispersion for each group, we take an average across groups weighting them with their share in the distribution of workers.

The table shows that for the benchmark, the dispersion decreases from 0.075 to 0.042 once controlling for the firm size (or firm productivity) variation. In other words, the heterogeneity in workers' employers explains 44 percent of the variation in hourly earnings in our benchmark model. The remaining variation is due to the hours variation within firms. In particular, the dispersion within hours bins in each firm productivity level is only 0.004, showing that the within firm size group dispersion in wages is entirely accounted for by the variation in hours.[17] Results also highlight that a crude grouping across hours worked does not reduce the dispersion in wages as much, because different size groups of firms exhibit different penalties for long and short and hours, and the reference usual hours in each firm differs with the firm size.

We contrast this computation with those coming out of the counterfactual exercises with a low degree of complementarities between hours of workers. Since the complementarities is the channel through which within-firm variation in hours translates into variation in wages, the model features no wage dispersion within firm size groups once we shut down these complementarities. We show this in the second row, with the counterfactual exercise of $\rho = 0.95$. In fact, the overall wage dispersion in the economy is lower than in the benchmark,

[17]The within size-hours bin dispersion would converge to zero as we narrow down the hours bins.

which reflects the lack of within firm wage dispersion.

Next, we turn our attention from the wage inequality to the inequality in total earnings (i.e., hourly wages multiplied by the hours worked). Naturally, factors that contribute to the hourly wage dispersion, as highlighted discussing the upper panel, also affect the total earnings dispersion. However, our model presents an additional mechanism through which the heterogeneities in firm TFP and worker hours jointly affect the overall earnings inequality. The sorting of long-hour workers into larger firms with higher TFP and hourly wages, amplifies the earnings differentials between workers. In the lower panel of Table 5, we show that controlling for working hours goes a long way in reducing the earning dispersion among workers, from a standard deviation in log-earnings of 0.29 to 0.06. This is not only because the hours increase total earnings for a given hourly wage, but also due to two of our model mechanisms. First, is the wage penalties related to the level of hours worked, as highlighted earlier. Second, is the sorting of workers with longer desired hours in to higher-pay firms. Since working hours systematically increase with the firm size, controlling for working hours means partially controlling for the size and the wage levels of the firms.

7 Final remarks

In this paper, we present empirical evidence from Canadian firms and workers that workers' average wage and average hours increase with establishment size, and the wage penalties for long (short) hours are larger in smaller (larger) establishments. We build a simple and novel theoretical framework to reconcile these facts. The model generates the size-wage premium, which endogenously translates in the workers' willing to work longer hours sort into larger firms. Through complementarities between workers' hours, this makes the long-hour wage penalties larger in smaller firms, and short-hour penalties larger in larger firms.

Our framework presents an environment to study the determinants of the wage inequality in the cross-section of workers. Quantitatively, we find that the within-firm dispersion in working hours is equally important as the firm-level heterogeneity in wages. In addition, we highlight a novel channel through which the sorting of workers with longer hours into larger firms, with higher hourly wage, contributes to the overall earnings dispersion.

References

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

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

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

Bloom, N., T. Kretschmer, and J. V. Reenen (2011): "Are familyâfriendly workplace practices a valuable firm resource?" *Strategic Management Journal*, 32, 343–367.

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.

Dvorkin, M., J. M. Sánchez, H. Sapriza, and E. Yurdagul (Forthcoming): "Sovereign Debt Restructurings," *American Economic Journal: Macroeconomics*.

Erosa, A., L. Fuster, G. Kambourov, and R. Rogerson (2017): "Hours, Occupations, and Gender Differences in Labor Market Outcomes," NBER Working Papers 23636, National Bureau of Economic Research, Inc.

Evans, D. S. and L. S. Leighton (1989): "Why Do Smaller Firms Pay Less?" *The Journal of Human Resources*, 24, 299–318.

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*.

Goldin, C. (2014): “A Grand Gender Convergence: Its Last Chapter,” *American Economic Review*, 104, 1–30.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): “Identifying Equilibrium Models of Labor Market Sorting,” *Econometrica*, 85, 29–65.

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

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

Lamadon, T., M. Mogstad, and B. Setzler (2019): “Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market,” NBER Working Papers 25954, National Bureau of Economic Research, Inc.

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.

Molina-Domene, M. (2018): “Specialization Matters in the Firm Size-Wage Gap,” CEP Discussion Papers dp1545, Centre for Economic Performance, LSE.

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.

Mueller, H. M., P. P. Ouimet, and E. Simintzi (2017): “Wage Inequality and Firm Growth,” *American Economic Review*, 107, 379–83.

Oi, W. (1983): “The Fixed Employment Costs of Specialized Labor,” in *The Measurement of Labor Cost*, National Bureau of Economic Research, Inc, 63–122.

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," *Handbook of labor economics*, 3, 2165–2214.

Papageorgiou, T. (2018): "Large firms and within firm occupational reallocation," *Journal of Economic Theory*, 174, 184–223.

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.

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–370.

Sohail, F. (2019): "Employer size and spinout dynamics," Working paper, Working Paper.

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.

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.

Appendix

A Evidence supporting taste shocks over firms of different productivity/size

Our model features two elements that are essential in replicating the motivating facts. We provide here an overview of the literature and empirical support for these features.

The taste shocks are introduced into the model to capture the data features that affect individuals' sorting into firms of different productivity and size, which are not accounted for by the current wages. In this section, we discuss several features of the labor market that motivate introducing the taste shocks in the model.

Non-pecuniary benefits. The taste shock captures, among other things, workers' heterogeneous preferences over non-pecuniary benefits (amenities). Jobs differ by their non-pay characteristics (Rosen, 1986), and recent studies such as Morchio and Moser (2018) and Lamadon et al. (2019) show the importance of workers' heterogeneous preferences over these non-pecuniary job characteristics in explaining the variation in earnings. One important dimension of such heterogeneity is gender. Compared with men, women desire more workplace flexibility and other non-pecuniary amenities, which helps explain a part of the gender pay gap (see Goldin, 2014, Morchio and Moser, 2018, and Erosa et al., 2017).

Importantly, we argue that these non-pecuniary characteristics might differ across firm size groups. Agell (2004) documented that managers in small firms rely less on pecuniary incentives to motivate their workers. They are also significantly more hostile towards incentive schemes based on competition and relative awards. The author argues that this is potentially consistent with workers of different types sorting into small versus large firms. For example, workers that value equality and friendly relations with their coworkers would prefer to work in small firms. Idson (1990) showed that workers in large establishments have a lower job satisfaction rate than those in small establishments, which could be explained mainly by greater rigidity in their work environment. Meanwhile, larger firms might dominate the smaller ones in some other dimensions such as safety in the work environment (Oi,

1974) and availability of female-friendly workplace practices (Bloom et al., 2011).

Pecuniary benefits. There are also other pecuniary benefits—other than hourly wages—that are important to workers. First of all, it is well-known that large firms provide more fringe benefits (Oi, 1983, and many others). Second, large firms seem to provide a better environment for learning and human capital accumulation. It is well-documented that large firms offer more training to their workers (Oi, 1983). Molina-Domene (2018) also documented that large firms assign workers to a specific task, whereas in small firms, workers often need to multi-task. As a result, workers in large firms are better able to develop task-specific human capital. Workers also learn entrepreneurial skills from their employers. Evidence has shown that previous employer size has a significant and positive correlation with the quality of the spin-out businesses (Sohail, 2019). Third, larger firms also have more occupations than small firms, making it easier for workers to reallocate to better-matched occupations through this internal labor market Papageorgiou, 2018). Lastly, since small firms have more variable growth rates and fail more often, jobs at smaller firms on average have a shorter duration than large firms, making small firms less attractive for workers who value stability. On the other hand, workers with a higher propensity to separate voluntarily or involuntarily are more likely to work for small firms (Evans and Leighton, 1989). This is consistent with the fact that younger workers ($<$25 years old) and older workers ($>$65 years old) are more likely to work for small firms (Headd, 2000).

Assortative matching. Workers are heterogeneous in ability and traits, making some workers more compatible with firms of certain size groups. Shimer and Smith (2000) provided a theoretical framework to study assortative matching between higher-ability workers and productive firms as a result of production complementarity. Empirically, Hagedorn et al. (2017) showed a robust positive assortative matching between workers and firms with a rank correlation of 0.75. They also estimated that, while observable worker and firm characteristics account for approximately 30 percent of the observed variation in wages, assortative matching can explain much of the remaining variation. The degree of assortative matching has also been increasing in the US, accounting for about one-third of the rise in inequality (Song et al., 2019). Since our model abstracts from heterogeneity in worker ability, we introduce taste shocks as a reduced-form way to capture the assortative matching between worker ability and firm productivity.

Labor market frictions. Firm size and productivity can be associated with other firm characteristics such as industry and location. Accordingly, there can be economic or logistical burdens forcing workers to work for specific productivity or size groups of firms in the market. Several studies have documented an urban-rural productivity gap, with urban firms more productive and larger than rural firms (Headd, 2000 and Melo et al., 2009). If workers are tied to a specific location, their moving cost or commuting cost contributes to the taste shocks we build into the framework. Headd (2000) also showed a cross-sector difference in firm size, with small 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. In this regard, the taste shock could be interpreted as the limited transferability of sector-specific human capital.

Overall, the discussion and the survey here support the appropriateness of the taste shocks in the context of self-selection into firms of different sizes. First, small and large firms differ in the multi-dimensional benefits (non-pecuniary or pecuniary) they offer to workers. Workers differ in the value they attribute to each of these dimensions. As a result, workers have heterogeneous preferences over the same type of firms. Second, the assortative matching between higher-ability workers and productive firms is an important feature of the labor market and contributes to wage inequality. We introduce the taste shocks to the model as a reduced-form way to capture this feature. Third, there can be logistical and technological reasons why different workers do not find it equally feasible to work in larger firms.

The taste shocks as we introduce them in our model are admittedly imperfect and reduced-form. First, most 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 these preferences. Second, our taste shocks are over firms of different productivity levels, even though the discussion above corresponds to the preferences or ability to work in firms of different sizes. In our model, size and firm-productivity will be isomorphic; hence, this language abuse would be innocuous in our framework.

B Technical relevance of ϵ-shocks

Our model features shocks to the values obtained by each occupation, which we describe in Section 4. The computational advantage of the ϵ-shocks is that they help convexify the occupational choice of workers by introducing additional randomness in their decision.

By assuming 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 occupations:

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

As mentioned in Section 4, having this probability as the policy function, instead of an indicator of 0 or 1 for choosing each occupation, smooths out the value function of workers, and help with convergence.

C Evidence from the US

In this section, we use micro-data from the US to replicate many of the findings from the Canadian LFS.

Data Description Our analysis uses pooled data from the 1995 to 2008 Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) which is conducted annually in March.[18] We start our sample in 1995 as it is the first year after a major redesign of the CPS in 1994 and end in 2008 to avoid confounding our findings with the impact of the Great Recession although we confirm that our main findings hold up to the latest available survey in 2019.

The ASEC contains detailed information on the economic activity of respondents from the previous calendar year. Important for our analysis are the variables which ask workers their earnings, usual hours worked and firm size. Firm size is reported in bins and records the total number of employees that a worker's employer has at all establishments.[19] We restrict attention to males aged between 25 and 64, who worked with a single private employer for 52 weeks in the previous year. We exclude those who usually worked fewer than 10 hours per week and earned less than half the federal minimum wage. and respondents with imputed values for firm size, hours or weeks worked. Our final sample includes around 267 thousand respondents.

<u>**Fact 1**</u> **Average Wages Increase in Firm Size** The wage premium in large firms has been extensively studied in the literature, see for example Oi and Idson (1999). Figure C.1 plots both the unconditional and conditional average log hourly earnings by hours worked and firm size. The figure establishes the firms size wage premium across three broad working hours categories in our sample. Table C.1 in the appendix shows that this result is robust to controlling for observable individual characteristics such as industry and occupation fixed effects.

[18] The underlying data-sets are extracted from IPUMS as detailed in Flood et al. (2020).

[19] For example the 2019 ASEC asks the following:
"Counting all locations where this employer operates, what is the total number of persons who work for your employer?"

Table C.1: Elasticity of Firm Size and Hourly Earnings

	All hours			$\leq$ 35 hrs	36-49 hrs	50+ hrs
Firm Size (Employees)	(1)	(2)	(3)	(4)	(5)	(6)
10 to 24	0.042***	0.033***	0.030***	0.038***	0.024***	0.039***
	(0.002)	(0.002)	(0.002)	(0.008)	(0.002)	(0.003)
25 to 99	0.075***	0.057***	0.051***	0.042***	0.046***	0.060***
	(0.001)	(0.001)	(0.001)	(0.008)	(0.002)	(0.003)
100 to 499	0.095***	0.074***	0.069***	0.056***	0.064***	0.078***
	(0.001)	(0.001)	(0.001)	(0.008)	(0.002)	(0.003)
500 to 999	0.106***	0.086***	0.082***	0.065***	0.076***	0.092***
	(0.002)	(0.002)	(0.002)	(0.011)	(0.002)	(0.003)
1000+	0.127***	0.105***	0.098***	0.060***	0.095***	0.104***
	(0.001)	(0.001)	(0.001)	(0.007)	(0.002)	(0.003)
Demographic Controls	Y	Y	Y	Y	Y	Y
Year, State FE	Y	Y	Y	Y	Y	Y
2-digit Industry FE	N	Y	Y	Y	Y	Y
2-digit Occupation FE	N	N	Y	Y	Y	Y
N	266,618	266,618	266,618	11,074	185,626	69,918
R^2	0.306	0.389	0.460	0.439	0.458	0.477

Notes: The table reports the estimated coefficient β_f from the following OLS regression:

$$\log(w_i) = \alpha + \sum_f \beta_f \mathbb{I}_{fi} + \gamma X_i + \epsilon_i$$

where w_i is the hourly wage of an individual $i.X_i$ is a vector of individual level controls which include a quadratic in years of experience, 4 education bins, race, marital status as well as controls for state, year, industry and occupation fixed effects. The variable $\mathbb{I}_{fi}$ is an indicator which is equal to 1 if i is employed in a firm of size f. The reference firm size category is firms with under 10 employees. Columns (4), (5) and (6) perform the regression for workers in different usual hours categories.

(a) Average Unconditional Wage (b) Average Conditional Wage

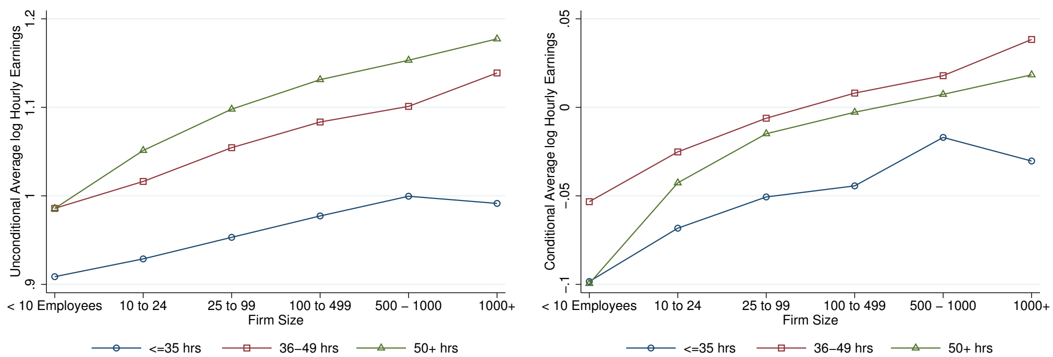

Figure C.1: The Firm-Size Premia

Note: Figure uses data from the pooled CPS sample. Panel (a) reports the unconditional average of log hourly earnings of workers by firm size and across hour bins. Panel (b) plots the average of conditional log hourly earnings. Conditional wages are the residuals from an OLS regression of log hourly wage on 4 education bins, a quadratic in years of experience, race and marital status dummies as well as year, state, 2-digit industry and 2-digit occupation fixed effects.

Fact 2 Average hours increase in Firm Size Focusing now on the relationship between firm size and average worker hours, Figure C.2 shows the distribution of hours worked by firm size. For clarity, we bin firm size into 3 categories; i) small (under 10 employees) ii) medium (between 10 and 100 employees) and iii) large firms (over 100 employees).

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

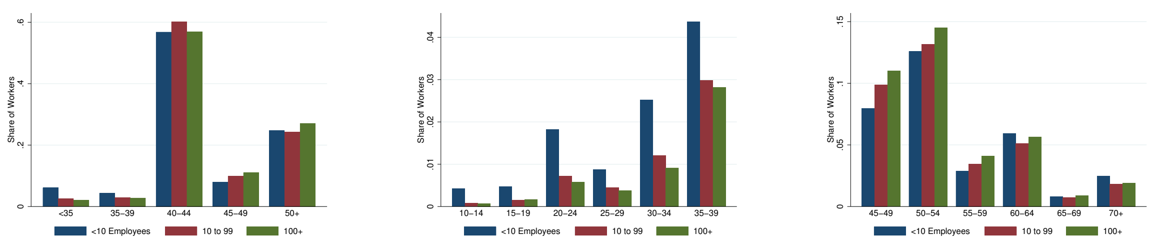

Figure C.2: Distribution of Working Hours by Firm Size

Note: Figure uses data from the pooled CPS sample.

Panel (a) shows the overall distribution of usual hours worked. While the median hours across all firms is 40-44 hours, there are important differences in the share of short and long hours worked by firm size. This can be seen in panel (b) and (c) which report the distribution of right and left tails of the hours distribution. From panel (b) it is evidence that workers in small firms are much more likely to work shorter (<39) hours than their counterparts in larger firms; while only around 4% of medium and large firm employees work 30-39 hours, the analogous share in small firms is around 7%. Panel (c) shows that medium and large firm employees are more likely to work between 45-59 hours but less likely to work very long (>60) hours. Indeed, the share of workers working greater than 60 hours is 9% in small firms compared to 7.7 and 8.5% in medium and large firms.

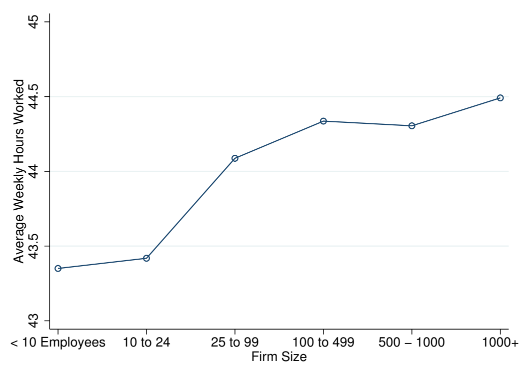

Figure C.3: Average Hours by Firm Size

Note: Figure uses data from the pooled CPS sample and reports the average weekly hours worked by firm size categories.

Figure C.3 plots average hours worked by firm size and shows that on average workers in the largest firms work 1 hour longer than workers in the smallest firms. While suggestive, the cross-sectional average does not control for confounding factors that might impact both firm size and hours worked. For example, certain industries might have a larger optimal scale and require long hours. Table C.2 reports the results from a regression on hours worked and a number of important individual worker level controls which include a quadratic in years of experience, 4 education bins, race, marital status as well as controls for state, year, industry and occupation fixed effects. The table shows that relative to firms with fewer than 10 employees, workers in firms of size 100+ work around 1.5 to 2% longer hours.

Fact 3 Long Hours Penalty Decreases in Firm Size Next, we study the cross-sectional relationship between wages and hours worked by firm size. To study the relationship between wages and hours worked, we estimated the following regression:

$$\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i$$

where w_i is the log hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works

Table C.2: Elasticity of Firm Size and Hours Worked

Firm Size (Employees)	(1)	(2)	(3)
10 to 24	0.007***	0.007***	0.008***
	(0.002)	(0.002)	(0.002)
25 to 99	0.018***	0.018***	0.017***
	(0.002)	(0.002)	(0.002)
100 to 499	0.021***	0.021***	0.020***
	(0.002)	(0.002)	(0.002)
500 to 999	0.016***	0.015***	0.017***
	(0.002)	(0.002)	(0.002)
1000+	0.016***	0.014***	0.016***
	(0.001)	(0.001)	(0.001)
Demographic Controls	Y	Y	Y
Year, State FE	Y	Y	Y
2-digit Industry FE	N	Y	Y
2-digit Occupation FE	N	N	Y
N	266,618	266,618	266,618
R^2	0.053	0.062	0.089

Notes: The table reports the estimated coefficient β_f from the following OLS regression:

$$\log(h_i) = \alpha + \sum_f \beta_f \mathbb{I}_{fi} + \gamma X_i + \epsilon_i$$

where h_i is the usual hours worked by an individual $i.X_i$ is a vector of individual level controls which include a quadratic in years of experience, 4 education bins, race, marital status as well as controls for state, year, industry and occupation fixed effects. The variable $\mathbb{I}_{fi}$ is an indicator which is equal to 1 if i is employed in a firm of size f. The reference firm size category is firms with under 10 employees.

h hours with the reference group being those that work between 40 and 44 hours. Control variables captured in X_i include a quadratic in years of experience, 4 education bins, race, marital status as well as controls for state, year, 2-digit industry and occupation fixed effects.

Figure C.4 plots the coefficient β_h for different firm sizes. First, we confirm the hump-shaped pattern in the wage-hours profiled previously documented in Yurdagul (2017) and Bick et al. (2020). That is, relative to working around 40 hours, there is a penalty for working either shorter and longer hours. By dis-aggregating across firm sizes, we also find that the long hours penalty is higher for small firm than for large firms. Indeed, working 70 to 74 hours in a small firm leads to 15% penalty relative to a 10% penalty in medium or large firms. Similarly, there is a 2.5% penalty of working 50 to 54 hours in small firms compared to no penalty in larger firms suggesting that the long hours penalty declines in firm size. There also appears to be some evidence of a lower short hours penalty in small firms but the evidence for this is statistically insignificant other than for those working between 30 and 34 hours.

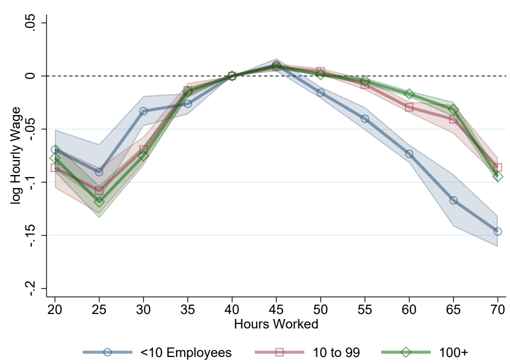

Figure C.4: The Relationship between Wages and Hours by Firm Size

Note: Figure uses data from the pooled CPS sample and plots the coefficient from a regression of log hourly earnings on hours worked dummies and demographic controls. The regression equation is $\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i$ where w_i is the hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works h hours with the reference group being those that work between 40 and 44 hours. The shaded regions are the 95% confidence interval.

D Additional Tables and Figures

Table D.3: Employment Share by Establishment and Firm Size

	Establishment	Firm
Less than 20	0.357	0.226
20 to 99	0.326	0.193
100 to 500	0.219	0.157
More than 500	0.098	0.424

Notes: The table reports the share of employees in the Canadian LFS in each Establishment and firm size category.

Table D.4: Accounting for earnings dispersion, LFS Data

	Overall	Within size-group	Within hours bin	Within size-hours bin
Wages	0.1176	0.1146	0.1158	0.1131
		(0.025)	(0.016)	(0.039)
Total Earnings	0.4356	0.4260	0.3450	0.3367
		(0.022)	(0.208)	(0.227)

Notes: The table reports the standard deviation of residual log wages and log total earnings from the pooled Canadian LFS data. Overall dispersion is the standard deviation of the residuals computed from a regression on log wages (or total earnings) as the dependent variable and the following controls; dummies for gender, marital status, controls for age, 2 digit industry and occupation as well as year and state fixed effects. The remaining columns report the standard deviation of residual log wages and earnings when also controlling for establishment size (column 2), hours worked (column 3) and both size and hours (column 4). Weekly hours bins are $\{20-, 20-22, ..., 72-74, 74+\}$. The terms in the parentheses report the share of overall dispersion accounted for by hours and size controls.

Table D.5: Accounting for Earnings Dispersion, LFS Data

	Hourly Earnings		Total Earnings	
	R^2	S.D.	R^2	S.D.
No Controls	0.0000	0.1574	0.0000	0.5866
Year, Province Controls	0.0378	0.1525	0.0266	0.5756
Age	0.0516	0.1515	0.0400	0.5718
Marital Status and Gender	0.1345	0.1428	0.1699	0.5219
Education	0.2092	0.1370	0.2245	0.5071
Industry	0.3133	0.1274	0.3246	0.4722
Occupation	0.4265	0.1176	0.4353	0.4356
Establishment Size	0.4555	0.1146	0.4611	0.4260
Hours (excl. establishment size)	0.4332	0.1158	0.5641	0.3450
Hours and Establishment Size	0.4593	0.1131	0.5846	0.3367
Hours, Establishment Size and Interaction	0.4599	0.1130	0.5851	0.3365

Notes: The tables reports the R^2 and dispersion of residual log earnings (either hourly or total) resulting from regressions with various controls. The first row reports these measures when there are no controls and each subsequent row includes additional controls. The last row reports these measures from a regression on log wages (or total earnings) as the dependent variable and the following controls; dummies for gender, marital status, controls for age, 2 digit industry and occupation as well as year and state fixed effects, establishment size, weekly hours worked and an interaction in size and hours. Weekly hours bins are $\{20-, 20-22, ..., 72-74, 74+\}$. The terms in the parentheses report the share of overall dispersion accounted for by hours and size controls.

Table D.6: Elasticity of Establishment Size and Hourly Earnings

	All hours			< 37 hrs	37 to 47 hrs	> 47 hrs
Establishment Size	(1)	(2)	(3)	(4)	(5)	(6)
20 to 99 Employees	0.029***	0.031***	0.030***	0.031***	0.028***	0.034***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
100 to 500 Employees	0.060***	0.056***	0.056***	0.059***	0.054***	0.054***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
More than 500 Employees	0.110***	0.096***	0.092***	0.094***	0.089***	0.089***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Demographic Controls	Y	Y	Y	Y	Y	Y
Year, Province FE	Y	Y	Y	Y	Y	Y
2-digit Industry FE	N	Y	Y	Y	Y	Y
2-digit Occupation FE	N	N	Y	Y	Y	Y
N	6,228,489	6,228,489	5,917,665	1,394,271	4,058,115	465,279
R^2	0.288	0.377	0.474	0.477	0.450	0.366

Notes: The table reports the estimated coefficient β_e from the following OLS regression:

$$\log(w_i) = \alpha + \sum_e \beta_e \mathbb{I}_{ei} + \gamma X_i + \epsilon_i$$

where w_i is the hourly wage of an individual $i.X_i$ is a vector of individual level controls which include a quadratic in years of experience, 4 education bins, race, marital status as well as controls for province, year, industry and occupation fixed effects. The variable $\mathbb{I}_{ei}$ is an indicator which is equal to 1 if i is employed in an establishment of size e. The reference size category is establishments with under 20 employees. Columns (4), (5) and (6) perform the regression for workers in different usual weekly hours worked categories.

Table D.7: Elasticity of Establishment Size and Hours Worked, without Gender/Marital Controls

Establishment Size (Employees)	(1)	(2)	(3)
20 to 99	0.046***	0.037***	0.034***
	(0.000)	(0.000)	(0.000)
100 to 500	0.066***	0.038***	0.036***
	(0.000)	(0.000)	(0.000)
More than 500	0.092***	0.039***	0.035***
	(0.000)	(0.000)	(0.000)
Demographic Controls	Y	Y	Y
Year, Province FE	Y	Y	Y
2-digit Industry FE	N	Y	Y
2-digit Occupation FE	N	N	Y
N	6,228,489	6,228,489	5,917,665
R^2	0.026	0.111	0.158

Notes: The table reports the estimated coefficient β_e from the following OLS regression:

$$\log(h_i) = \alpha + \sum_e \beta_e \mathbb{I}_{ei} + \gamma X_i + \epsilon_i$$

where h_i is the usual hours worked by an individual $i.X_i$ is a vector of individual level controls which include a quadratic in age, 4 education bins as well as controls for province, year, industry and occupation fixed effects. The variable $\mathbb{I}_{ei}$ is an indicator which is equal to 1 if i is employed in an establishment of size e. The reference establishment size category is establishments with under 20 employees.

(a) No Industry or Occupation Controls

(b) Both Industry and Occupation Controls

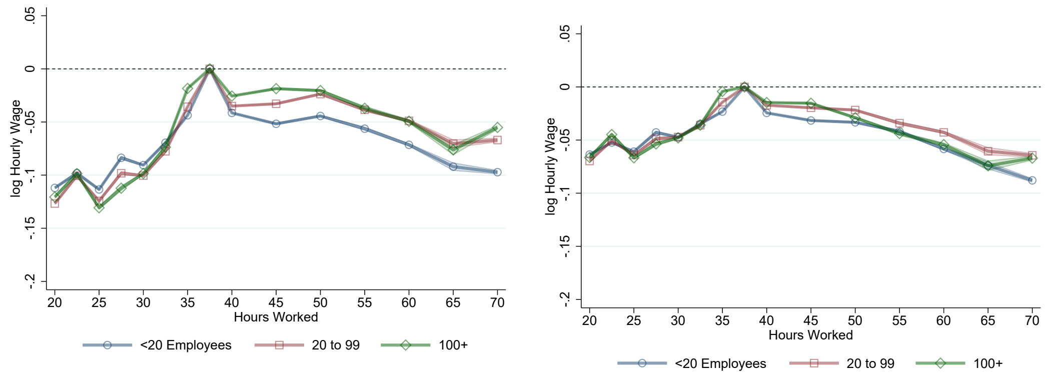

Figure D.5: The Relationship between Wages and Hours by Establishment Size

Note: Figure uses data from the pooled CPS sample and plots the coefficient from a regression of log hourly earnings on hours worked dummies and demographic controls. The regression equation is $\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i$ where w_i is the hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works h hours with the reference group being those that work between 40 and 44 hours. The shaded regions are the 95% confidence interval.

Figure D.6: The Relationship between Wages and Hours by Establishment Size, no gender/Marital Status

Note: Figure uses data from the pooled LFS sample and plots the coefficient from a regression of log hourly earnings on hours worked dummies and demographic controls. The regression equation is $\log(w_i) = \alpha + \sum_h \beta_h \mathbb{I}_{hi} + \gamma X_i + \epsilon_i$ where w_i is the hourly wage and $\mathbb{I}_{hi}$ is an indicator equal to 1 if individual i works h hours with the reference group being those that work at least 37.5 hours and less than 40 hours. The shaded regions are the 95% confidence intervals.