No. 38

JUNIO DE 2017

Documentos CEDE

ISSN 1657-7191 Edición electrónica.

Managing skill premium through
firm structure

Enoch Hill

David Perez-Reyna

CEDE
CENTRO DE ESTUDIOS SOBRE DESARROLLO ECONÓMICO

Universidad de
los Andes
Facultad de Economía

Serie Documentos Cede, 2017-38
ISSN 1657-7191 Edición electrónica.
Junio de 2017

© 2017, Universidad de los Andes, Facultad de Economía,
CEDE. Calle 19A No. 1 – 37 Este, Bloque W.
Bogotá, D. C., Colombia Teléfonos: 3394949- 3394999,
extensiones 2400, 2049, G 11
infocede@uniandes.edu.co
<http://economia.uniandes.edu.co>

Impreso en Colombia – Printed in Colombia

La serie de Documentos de Trabajo CEDE se circula con propósitos de discusión y divulgación. Los artículos no han sido evaluados por pares ni sujetos a ningún tipo de evaluación formal por parte del equipo de trabajo del CEDE.

El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y sólo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor, sólo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair use), estén previa y expresamente establecidas, no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular, y no atenten contra la normal explotación de la obra.

Universidad de los Andes | Vigilada Mineducación
Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

Managing skill premium through firm structure*

Enoch Hill[†]

David Perez-Reyna[‡]

Abstract

Previous research has consistently demonstrated a positive relation between firm size and skill premium. We decompose this result by type of skilled worker using data from Chilean firms and find that returns in skill premium to size are an order of magnitude larger for owners and managers compared to other types of skilled workers. We interpret this to imply that firm structure is important for understanding the relationship between skill premium and firm size. We propose an original model of firm hierarchy which rationalizes our empirical finding. In our model, more productive firms maximize profits by choosing to employ more levels of management. These managers require a higher average skill level and command a higher average skill premium. Additionally, our model predicts that skill premium is increasing in the ratio of workers to managers, a fact we also observe in the Chilean data.

JEL Classification: J31, L22, D23

Keywords: Skill premium, firm structure, human capital

*We would like to thank James A. Schmitz, Jr., Timothy Kehoe, Cristina Arellano, Fabrizio Perri, Terry Roe and Carlos Medina for their valuable comments. We are particularly indebted to Michael Maio as this work builds off of earlier joint work (Maio, 2014). We are also grateful to the participants at the Workshop in International Trade and Development at the University of Minnesota and seminar participants at the University of Exeter, Universidad de los Andes and the IMF, as well as conference participants at Segundo Network de Microeconomía - AFADECO.

[†]e-mail: enoch.hill@wheaton.edu, Wheaton College

[‡]e-mail: d.perez-reyna@uniandes.edu.co. Universidad de los Andes

Administrando la prima por habilidad por medio de la estructura de las empresas

Enoch Hill David Perez-Reyna

Resumen

La literatura ha encontrado una relación positiva entre el tamaño de una empresa y la prima por habilidad. En este artículo descomponemos este resultado por tipo de trabajador habilidoso usando datos de empresas chilenas. Encontramos que los retornos en la prima de habilidad relativo al tamaño de una empresa son mayores en un orden de magnitud para dueños y administradores de empresas, comparado con otros tipos de trabajadores habilidosos. Esto implica que analizar la estructura de una empresa es importante para entender la relación entre prima por habilidad y el tamaño de una empresa. Proponemos un modelo de jerarquía en las empresas que racionaliza este resultado empírico. En nuestro modelo las empresas más productivas maximizan beneficios por medio de la escogencia de más niveles de administración. Los administradores en estos niveles requieren una mayor habilidad y demandan una mayor prima por habilidad. Nuestro modelo predice que la prima por habilidad es creciente en la razón entre administradores y trabajadores, que es consistente con lo que observamos en los datos chilenos.

Clasificación JEL: J31, L22, D23

Palabras clave: prima por habilidad, estructura de las empresas, capital humano

1 Introduction

In this paper we bring new data and new theory to bear on a longstanding question: Why do large firms pay higher wages than small firms? While many explanations focus on the fact that large firms employ disproportionately more labor-augmenting technology than smaller firms (e.g. Dunne and Schmitz (1995), Dunne et al. (2002)), we propose a different story: Larger firms choose organizational structures with more layers of management, and the wages paid to managers increase as the number of layers grows.¹ We develop a model in which managers near the top of the management hierarchy in large firms provide a larger marginal contribution to output than their counterparts at smaller firms since they oversee a larger quantity of firm activities. As a result, they receive higher wages.

Our model is consistent with empirical evidence that the wage premium offered by large firms is greater among skilled employees than among unskilled employees (e.g. Brown and Medoff (1989), Davis et al. (1991), and Idson and Oi (1999)). Equivalently, the skill premium (the ratio of the average wage a firm pays to skilled employees to the average wage it pays to unskilled employees) is increasing in firm size.

We are able to go a step further, using recent data from Chile’s manufacturing sector to show that managers in particular receive a large wage premium at large firms - much more so than other categories of skilled workers. The empirical literature typically uses broad definitions of skilled and unskilled labor to study the skill premium. For example, some studies (e.g. Dunne and Schmitz (1995) and Pavcnik (2003)) treat blue-collar workers as skilled and white-collar workers as unskilled even though there are substantial skill differences within each group. That is, both a CEO and a data entry employee would be classified as white-collar. By using detailed plant-level data on employment and wages from Chile’s Manufacturing Industry National Survey (ENIA) during the period 2000-2007, we are able to obtain a more granular picture of the wage premia offered by large firms within different labor categories. We construct and study several alternative measures of the skill premium at the firm level. One measure is the ratio of the average wage a firm’s managers receive to the ratio of wages a firm’s unskilled production workers receive. Another is the ratio of average wages received by skilled production workers (“technicians”) to average wages received by unskilled production workers. The latter ought to most directly capture any skill-bias in the technology hired by large firms.

What we find is striking. The effect of firm size on the manager skill premium is 9 to 14 times greater than the effect of size on the skill premium of technicians or administrative

¹It should be noted that these stories are not necessarily at odds with one another. It is feasible that the labor-augmenting technology augments the capacity of managers and in this way, increases both productivity and wages.

personnel (other white-collar workers). Our results suggest that managerial inputs and firm hierarchies are an important reason why large firms pay a higher skill premium than small firms. This is consistent with the work of Nikolowa (2010) who suggests that skill biased *organizational* change contributes to observed increases in skill premium. We make two additional observations from the Chilean data. First, as firms grow, the ratio of workers to managers tends to increase. We call this an increase in the average “width” of larger firms as each manager is responsible for more workers. Second, skill premium is increasing in both the width of firms and the size of firms.

To investigate the relationship between firm size and the skill premium theoretically, we develop a hierarchical model of firms that borrows elements from Garicano (2000) and Caliendo and Rossi-Hansberg (2012), where large firms optimally choose organizational structures with more layers of management than small firms. A manager receives direct reports from employees on the next-lowest layer (either production workers or lower-level managers) and is able to augment their output multiplicatively using her managerial labor. High-level managers at large firms, who have the scope to augment large amounts of output, receive high wages. Small firms, producing relatively little output, have no such high-productivity managers and therefore pay their managers lower wages on average. Additionally, since wages for unskilled workers are constant across firms in our model, the skill premium is increasing in firm size. These results can be obtained as the equilibrium outcome of a general equilibrium outcome.

Traditional models of human capital are indifferent between hiring many low skilled workers or fewer high skilled workers. Our model has the added benefit of explaining why firms choose differing compositions of high and low skilled labor. Through the lens of our model, the reason bigger firms pay a higher skill premium is that they demand a higher average level of human capital in their employees relative to their smaller firm counterparts.

One novel feature of our model is that the number of layers of management a firm chooses is a continuous, rather than discrete variable. Doing so gives the model substantial flexibility and allows us to map the firm’s production function into a class of familiar production functions in which workers differ only in the units of effective labor they supply. The advantage of our model, relative to those with just human capital, is that it allows us to evaluate firm hierarchical choices. While the notion of having, say, 2.5 layers of management is not immediately intuitive, we attempt to provide some clarity when the model is presented.

This paper is connected to an extensive empirical literature on the relationship between firm size and wages. Brown and Medoff (1989) provided an early contribution by showing that large firms tend to pay their workers higher wages than smaller firms. The same phenomenon has been documented by many other researchers. For instance, Davis et al.

(1991) study U.S. Census manufacturing data from 1963-86 and find a steep monotonic relationship between a firm's size (in terms of number of employees) and the average wage paid to its workers. Idson and Oi (1999) corroborate their results using more recent Census data.

Most empirical work, including the Davis et al. (1991) study cited above, also shows that the wage gap between large and small firms is the largest among skilled workers, indicating that the within-firm skill premium increases, on average, with firm size. Haskel (1998) provides a counterpoint by showing that, among UK manufacturing firms in the 1980s, small firms had a higher skill premium than larger firms.

Many explanations have been advanced for the wage premium paid by large firms (for a summary, see Katz and Summers (1989)). One explanation for why large firms pay higher wages than small firms that has garnered substantial empirical support is that large firms demand more skilled labor than small firms. For example, Abowd et al. (1999) provide evidence from France, and Zweimüller and Winter-Ebmer (2003) provide evidence from Switzerland. One reason why large firms demand more skilled workers is that they need skilled labor to operate more advanced production technology (See Dunne and Schmitz (1995) or Dunne et al. (2002)). While differences in technology appear to contribute meaningfully to the wage gap between large and small firms, a substantial portion of the variation remains unexplained by technological differences alone.

The technological explanation for the firm-size wage gap has been adopted by researchers seeking to understand the effects of international trade on labor markets. Bernard and Jensen (1999) document that exporting firms, which tend to be the largest firms, raise the demand for skilled labor relative to unskilled labor, leading to increases in the skill premium. Yeaple (2005) and Bustos (2011) present models explaining exporters' high demand for skilled labor. By serving a large international market, exporters have a stronger incentive to invest in cost-reducing technologies than firms that only serve a small local market. To the extent that the new, labor-saving technologies require skilled labor to operate, large exporting firms will hire more skilled labor than smaller firms. Helpman et al. (2010) further develop this argument theoretically.

Our theoretical approach to studying the firm-size wage gap differs from existing theories in that we focus on the effect of firm size on organizational structure. While our model is consistent with the fact that larger firms demand higher-skilled employees than smaller firms, we develop explicitly the idea that larger firms choose a hierarchical structure with more levels than smaller firms. To do so, we draw on the hierarchical model of Garicano (2000) in which firms optimally choose the number of levels of management by trading off gains from having more managers solve problems against the costs of communication between many

levels of management. Various papers have built on this structure: Beggs (2001), Caliendo and Rossi-Hansberg (2012), Garicano and Rossi-Hansberg (2006), Garicano and Hubbard (2007), Grobovšek (2013) and Garicano and Hubbard (2009), among others. The fact that firms organize in hierarchical levels, and that these levels are knowledge-based have been corroborated in data by Bloom et al. (2014) and Caliendo et al. (2015).

Relative to their models, our main theoretical contribution is to model the number of management levels that a firm chooses as a continuous, rather than discrete, choice. Doing so demonstrates that our model is essentially a standard, simple human capital model with a restriction on firms which require them to hire specific ratios of workers across levels of ability. This provides explicit predictions about firm organizational structure and skill premia. Instead of labor being perfectly substitutable in production, imposing a structural hierarchy upon firms turns employees with varying levels of human capital into complements. Two workers cannot simply replace a manager but must work in conjunction with managers in order to accomplish additional tasks. Larger firms require more levels of management and a higher average level of human capital, and consequently they pay a higher average wage.

A question remains as to whether workers at bigger firms are paid more because they are more skilled, or because something about the firm structure augments their productivity. If it is the later, it is not clear why these workers are paid more, since it is the firm and not the worker that possesses the labor augmenting technology. We do currently lack any worker level data for Chile. Consequently, we are unable to demonstrate empirically that these higher levels of skill premium are connected to higher levels of human capital at larger firms.²

The paper proceeds as follows: In Section 2 we present our data findings. In Section 3 we introduce a firm structure to allow for a flexible selection of both levels and measure of employees. Section 4 presents and characterizes a general equilibrium model using the firm structure introduced in the previous section. Finally, Section 5 concludes.

2 Data

In this section we use data from the Manufacturing Industry National Survey (ENIA) compiled by the Chilean National Statistics Institute (INE) to analyze the relationship between size and skill premium. This database allows us to differentiate between expenditures on different types of skilled workers. In this way we are able to explain more precisely how size

²Fox (2009) addressed this issue by using US data and finds that the wage gap increases with responsibility, which is positively correlated with span of control. In our model managers need to get more education to manage more workers, so we don't have observationally equivalent consumers receiving different wages.

affects skill premium.

ENIA compiles economic and accounting information for registered firms that are established in Chile and have 10 or more employees.³ The survey includes annual information from 1995 to 2007 for around 4,600 firms per year with just over 8,000 firms entering some portion of the sample throughout the time period. Information differentiating types of skilled workers is available beginning in 2000.

The information allows us to distinguish between skilled and unskilled workers, as well as how much was spent on the wage bill for employees of each type. Unskilled workers are non specialized workers in charge of executing tasks, mainly manual, that are directly related to the production process (“Blue Collar”). We categorize all other employees as skilled workers (“White Collar”).⁴ Information regarding the average size of firms and the breakdown by worker type is included in Table 1.

We define the skill premium of a firm as the ratio of the average wage per skilled worker to the average wage per unskilled worker. We are able to identify different types of skilled workers in the data. In particular skilled workers can be further divided into owners, managers, technicians and administrative personnel. For each firm we estimate multiple measures of skill premium: a skill premium for each type of skilled worker and a skill premium aggregating across all skilled workers. While perhaps a slight misnomer, in the total row we also provide the ratio of the wage for the average worker at a firm (including unskilled workers) relative to the wage for the average unskilled worker. Table 1 shows the resulting skill premium by type of worker. Note that the average manager gets paid over six times the wage of an unskilled worker. This number is much higher than the skill premium for other types of skilled workers.⁵

Next, we analyze the impact of the size of a firm on the skill premium. To do this we run the following regression:

$$SP_{it} = \beta_0 + \beta_1 size_{it} + \alpha_t + \gamma_i + \varepsilon_{it}, \quad (1)$$

where $size_{it}$ is the natural logarithm of the total number of workers at firm i in period t , α_t is a time fixed effect, and γ_i is a firm fixed effect. Table 2 shows that the skill premium is increasing in the size of a firm. This result is robust across all types of skilled workers. Moreover, size seems to affect the skill premium of owners and managers more than it affects

³INE targets firms with 10 or more employees. If a firm has more than one establishment, each establishment is reported separately. In this case, there can be reports by establishments with under 10 employees.

⁴Only full-time workers are included in the analysis.

⁵Owners constitute a much larger fraction of full time employees at smaller firms. As skill premium is increasing in firm size, this helps explain much of the difference in skill premium between managers and owners.

Table 1: Summary Statistics by WorkerType

	avg. size	size std. dev.	SP avg.	SP wtd. avg.	SP std. dev.
Managers	1.92	4.68	6.05	7.20	4.54
Owners	0.73	1.16	4.39	3.28	4.08
Managers and Owners	2.64	4.75	5.37	6.12	4.21
Technicians	20.47	49.26	1.78	1.91	0.92
Administrative Personnel	8.45	17.68	1.48	1.76	0.72
All skilled workers	31.56	60.71	2.05	2.22	0.84
Unskilled workers	32.50	64.64	1.00	1.00	NA
Total	64.06	100.67	1.38	1.60	0.38

Includes data from 2000 to 2007. Weighted average skill premiums are larger than the simple average skill premiums since larger firms tend to pay larger skill premiums.

Source: INE. Authors' calculations.

the skill premium of other types of skilled workers. In particular, the effect of size on the skill premium of technicians is very small: Holding all other variables constant, an increase of 1% in the size of a firm raises the skill premium of technicians by 0.085%. That is, even though there is a positive relation, the size of a firm does not seem to be economically significant in determining the skill premium of technicians. The coefficient for owners and managers is more than 14 times larger than for technicians and 9 times larger than for administrative personnel.⁶ Each 1% increase in the size of a firm predicts more than a 1% larger skill premium.

The regression results suggest that bigger firms pay higher wages for skilled workers due to a higher marginal productivity of owners and managers, not of technicians or administrative personnel. In particular, if the theory that acquisition of skill biased capital can explain why larger firms offer a higher skill premium is true,⁷ the data suggests that this skill biased capital is particularly complementary to the marginal productivity of managers and owners as opposed to other categories of skilled labor. Regardless the view one takes on skill biased technical change, our data analysis supports the importance of considering the organizational structure of a firm when analyzing the relation between size and skill premium.

At first viewing of Table 2, it may appear odd that the coefficient on size for all skilled workers falls outside the range of any coefficient for the individual contingents of skilled workers. A deeper look suggests that this occurs due to a change in the composition of skilled workers by firm size. Larger firms tend to employ a smaller proportion of managers

⁶We get similar results if we calculate the skill premium using average hourly wage instead of average annual wage.

⁷As has been proposed in the literature (e.g. Dunne and Schmitz (1995) and Dunne et al. (2002)).

Table 2: Skill Premium vs Size of Firm

	Managers	Owners	Managers and Owners	Technicians	Administrative Personnel	All skilled workers
Size	1.261*** (0.120)	1.051*** (0.131)	1.231*** (0.087)	0.085*** (0.022)	0.136*** (0.018)	0.078*** (0.017)
Year fixed effects	yes	yes	yes	yes	yes	yes
Firm fixed effects	yes	yes	yes	yes	yes	yes
Number of Observations	12, 196	8, 480	17, 064	16, 687	18, 367	20, 418

OLS estimation. Standard errors are in parenthesis.

***: Significant at 1%.

Source: INE. Authors' calculations.

and owners relative to technicians and administrative personnel. Even though skill premium is increasing in firm size for all types of skilled workers, skill premium is largest for managers and owners for firms of any size. As firm size increases, the fraction of managers and owners falls. This exerts downward pressure on the coefficient between skill premium and size causing it to lie outside the range of coefficients for any individual component regression of skilled workers.

To understand how the composition of skilled workers changes across firm size, we define *MOfrac* as the number of managers and owners relative to all skilled workers and we run the following regression

$$MOfrac_{it} = \beta_0 + \beta_1 size_{it} + \alpha_t + \gamma_i + \varepsilon_{it}.$$

Regression results are included in Table 3. The coefficient on size suggests that a doubling in firm size decreases the fraction of managers and owners that comprise all skilled workers by 6.4%. If we assume these managers and owners are replaced with skilled workers who receive the average skill premium across all firms, the skilled worker composition effect lowers the estimate on the coefficient for size by roughly 0.212.⁸ Without the composition effect, we expect the relation between firm size and skill premium would be greater than or equal to 0.290. Our assumption that managers and owners would be replaced by skilled workers who receive the average skill premium across all firms underestimates the composition effect since these workers are necessarily replaced by technicians or administrative personnel who receive, on average, a skill premium smaller than the average skill premium of all skilled workers. This is an important result since it suggests that prior estimates of the effects of

⁸This number is calculated from $0.064 \times (5.37 - 2.05)$. 5.37 and 2.05 represent the simple average skill premium for managers and owners and all skilled workers respectively.

firm size on skill premium may be significantly underestimated if they wish to understand the effects of firm size on skill premiums *for a specific worker*. At the very least, in the Chilean data used here, the changes in composition of skilled workers across firms size masks the majority of the effects of firm size on skill premium for workers of a specific type.

Table 3: Skilled Worker Composition by Firm Size

Size	-0.064*** (0.003)
Year fixed effects	yes
Industry fixed effects	yes
Number of Observations	12, 196

OLS estimation. Standard errors are in parenthesis.

***: Significant at 1%.

Source: INE. Authors' calculations.

Next, we analyze the relationship between the organizational structure of a firm and the skill premium. For this we extend regression (1) to control for the ratio of unskilled workers to managers and owners. We label this ratio as span of control, SC . The extended regression is as follows:

$$SP_{it} = \beta_0 + \beta_1 size_{it} + \beta_2 SC_{it} + \alpha_t + \gamma_i + \varepsilon_{it},$$

where α_t and γ_i are defined as in (1). Table 4 shows the results. We find that the higher the number of unskilled workers per manager or owner, the larger the difference in wage. Note, the regression coefficient of 0.017 implies a 1.7% increase in wage for each additional direct report.⁹ This supports the idea that firms where managers have a wider span of control offer a higher compensation to managers.

These data findings motivate the structure of firms used in our model. We will focus on the role of managers and owners, abstracting from other types of skilled workers, since it is the wages of managers and owners that drive the relation between size and skill premium. Additionally, the firm structure should explain why managers and owners with more direct reports get paid more.

⁹The regression coefficient on size suggests that doubling the size of the firm leads to a 98% increase in skill premium or that a 1% increase in the size of the firm is related to a 0.98% increase in skill premium for managers.

Table 4: Skill Premium for Managers and Owners vs Composition of Firm

Size	0.977*** (0.089)
Span of control	0.017*** (0.001)
Year fixed effects	yes
Industry fixed effects	yes
Number of Observations	17,064

OLS estimation. Standard errors are in parenthesis.

***: Significant at 1%.

Source: INE. Authors' calculations.

3 Model

We first set up a discrete firm structure in order to introduce some intuition for our main results. Then we extend our firm structure to a continuous setting to derive our main comparative statistics.

3.1 Discrete firm structure

To develop insight into the relationship between firm size and the skill premium paid to managers, we develop a hierarchical model of firms that draws elements from Garicano (2000) and Caliendo and Rossi-Hansberg (2012). Briefly, a firm consists of several levels of employees, with employees on the bottom level representing production workers and employees on higher levels representing managers. We assume that within a firm the width of each level is constant, although this width could change across firms. This width is determined by how many direct reports managers receive from the level strictly below, which is precisely the span of control. Throughout, we will use interchangeably the terms width of a level, width of a firm, and span of control.

Production workers complete a number of tasks, but a fraction of their potential output is lost before their tasks are converted into the final good of the firm. Managers on the first level above production workers receive the uncompleted tasks from the workers who report to them. These managers are in turn able to complete a portion of the unfinished tasks. Tasks that even a manager on the first level cannot complete are handed to the next level of management, and so on. A firm optimally chooses the number of levels of management and the span of control to balance the benefits of being able to solve additional tasks against the costs of compensating additional managers.

In our full model, the number of management levels and the span of control chosen by a firm are continuous variables. We first develop intuition by presenting a simple version of the model in which both the number of management levels and the span of control are discrete.

Consider a firm with X levels and span of control equal to R . The structure of the firm is rigid in the sense that specific quantities of managers at each level are required. Each worker at level $x = 0$ receives a measure 1 of tasks. Of those tasks, the worker is able to solve a share equal to $1 - \lambda$, and the uncompleted fraction $\lambda \in (0, 1)$ of the tasks is passed to the level immediately above. No two employees work on the same task. Each manager in level $x > 0$ is assigned R direct reports from the level immediately below, so they receive a measure $R\lambda$ of unsolved tasks, of which they are able to solve a share $1 - \lambda$. The unsolved fraction λ of the tasks are passed in turn to the next-highest level of management. The quantity R is constant within a firm so that a manager above the first management level receives direct reports from R managers on the level immediately below. A firm's total output is the number of tasks its workers and managers are collectively able to solve.

We denote employees at the lowest level ($x = 0$), who receive no reports, as workers, while employees on higher levels are managers. We will denote managers at the highest level as CEO's and we assume that every firm has one CEO. Figure 1 displays the case where $R = 2$ for firms with $X = 1$ and $X = 2$ to clarify our idea of a firm.

Figure 1

(a) Firm with one level of management
($X = 1, R = 2$)

(b) Firm with two levels of management
($X = 2, R = 2$)

Each firm chooses the optimal number of levels of management, X , and the span of control, R . Hiring additional levels of management allows more tasks to be solved, and increasing span of control allows more tasks per manager, but both situations will result in a higher wage bill, as will be specified in the next section. Specifically, a firm with X levels hires R^{X-x} managers at level $x > 0$ and R^X workers. Furthermore, it is able to solve $R^X(1 - \lambda^{X+1})$ tasks. Lemma 1 in Appendix A states and proves this result.

Assume that wages of all employees - both workers and managers - are proportional to

the number of tasks they complete (as they will be in the general equilibrium version of the model). Notice that if $R\lambda > 1$ then the contribution of managers is increasing in the level of management x , so wages would be increasing in the contribution of managers. Otherwise the wages of managers would be lower than the wages of workers, which is at odds with data. Define the skill premium of a firm, $SP_{X,R}$, as the ratio of the average wage of managers to the average wage of workers. Then $SP_{X,R}$ can be expressed as

$$\begin{aligned} SP_{X,R} &= \frac{\sum_{x=1}^X R^{X-x} (R\lambda)^x (1-\lambda)}{\sum_{x=1}^X R^{X-x}} \times \frac{1}{1-\lambda} \\ &= \frac{\lambda}{1-\lambda} (R-1) \frac{R^X}{R^X-1} (1-\lambda^X), \end{aligned}$$

where the first term is the average contribution of managers and the second is the inverse of the average contribution of workers. We find that the skill premium is increasing in both levels of management and span of control as long as $R\lambda > 1$.¹⁰

The assumption $R\lambda > 1$ implies that each manager completes more tasks than each of her direct reports individually. The lemma establishes two results central to the theme of our paper for the simple, discrete version of the model. The skill premium increases in firm height (proxied by X) because firms with more levels have top managers who receive many reports (both direct and indirect) and are therefore able to complete many tasks. The second part of the lemma results from the fact that managers who receive more direct reports have the opportunity to solve more tasks that employees at lower levels were unable to complete. On average, then, managers at large firms receive higher wages than managers at small firms.

3.2 Continuous firm Structure

We now extend the firm structure to allow for a flexible (continuous) selection of both levels and measure of employees. Consider a firm with X levels and span of control R . To maintain consistency between the discrete version of the model and the continuous version, we assume a width of measure one of CEOs. The span of control R determines the width of managers at level x to be R^{X-x} . Let manager $n \in [0, R^{X-x} - 1]$ on level x be a manager who occupies a one-dimensional horizontal line with a width of measure one and a left edge equal to n . We define the “area of command” for manager n located on level $x > 0$ to be the set of workers and managers who either report directly to manager n or report to a manager within the area of command for manager n . Manager n is in charge of dealing with the unsolved tasks

¹⁰Lemma 2 in Appendix A states and proves this result formally.

of everyone within her area of command. Formally, the area of command of manager n on level $x > 0$, $AC(x, n)$, is defined as

$$AC(x, n) \equiv \left\{ (x', n') : x' \in [0, x), nR^{x-x'} < n' \leq (n+1)R^{x-x'} \right\}.$$

To map the concept of area of command to the discrete case, consider a firm with 2 levels and span of control $R = 3$. The area of command of manager $n = 1$ on level $x = 1$ is the three workers on the bottom level located strictly to the right of position 3 ($= nR^x$) and to the left of (including) position 6 ($= (n+1)R^x$). Figure 2 illustrates a sample area of command.

We find that there is a continuum of measure $\frac{1}{\ln R}(R^x - 1)$ of employees in area of command $AC(x, n)$ and, if there is a measure 1 of CEOs, there is a measure $\frac{R^X - 1}{\ln R}$ of managers in a firm with X levels and span of control R . (See Lemma 3 in Appendix A).

Notice that the measure of employees in an area of command doesn't depend on the location of manager n on level x . Therefore, without loss of generality, we'll denote areas of command and the tasks solved within them by the level of the manager. That is, the area of command of a manager in floor x will be denoted by $AC(x)$.

Figure 2: $AC(x)$

Similar to before, every employee is able to solve a measure $1 - \lambda$ of the tasks she is assigned. The remaining fraction is sent to a manager located in a higher level. As before,

the measure of tasks each employee is able to solve is independent of the tasks any other employee can solve. The number of reports that a manager in level x gets from employees in level $x - \Delta$ is equal to R^Δ . Intuitively, as we transition from the discrete to continuous case, levels come closer together and each manager gets less reports from the level directly below them, but there are more levels below them. Additionally, they are able to solve only a fraction $(1 - \lambda)\Delta$ of the tasks they receive. With these assumptions, the total number of tasks that are solved by the area of command of a manager working at level x in a firm with width R is

$$\mathcal{T}(x, R) = R^x (1 - e^{-x(1-\lambda)}) .$$

If there is a measure 1 of CEOs, then the total number of tasks solved by a firm with X levels and R reports is $R^X(1 - e^{-X(1-\lambda)})$, as is formally stated and proved in Proposition 1 in Appendix A.

The area of command of a manager in level x , is denoted $AC(x)$. This manager receives a measure R^x of tasks to be solved. Since the total measure of tasks solved within this area of command is $R^x (1 - e^{-x(1-\lambda)})$, there is a measure $R^x - R^x (1 - e^{-x(1-\lambda)}) = R^x e^{-x(1-\lambda)}$ of unsolved tasks that a manager at floor x faces. Of this measure she is able to solve a fraction $1 - \lambda$. Therefore the marginal contribution of a manager at level x , MC_x is given by

$$MC_x \equiv R^x e^{-x(1-\lambda)}(1 - \lambda). \quad (2)$$

In our model the structure of firms is rigid in the sense that a firm with X levels must employ R^{X-x} employees at each level x . In this sense, the firm faces a modified Leontief production function for which the proportions of employees hired at each level are fixed. The added element of firm structure converts employees at various levels into perfect complements. Consequently, the concept of marginal productivity of a specific employee is unclear since production occurs through the joint effort of all employees at the firm. This implies that labor remuneration in a competitive equilibrium setup relies heavily on the supply of managers and workers. It seems natural that labor should be remunerated according to the marginal number of tasks that workers and managers are able to solve. Therefore, for the rest of the paper we assume that the wage of workers is equal to $1 - \lambda$ (workers receive a measure 1 of tasks but are only able to solve a measure $1 - \lambda$), and the wage of a manager in level x with span of control R is MC_x . In Section 4 we provide a sample general equilibrium framework in which this assumption on wages is an endogenous outcome of the economy.

With this assumption, if $MC_x \leq 1$, managers would have a lower wage than workers. Similar to what we have in the discrete model, we don't consider this case since this would

be at odds with data. Since $MC_x > 1$ only when

$$\ln R > 1 - \lambda,$$

we restrict R to satisfy $R > e^{1-\lambda}$.

3.3 Comparative statistics

Recall that in Section 2 the primary findings from the Chilean data set were three-fold. First, managers (as opposed to technicians or administrative personnel) are the primary driver of skill premium. Second, skill premium is increasing in the size of the firm when controlling for span of control. Third, skill premium is increasing in the span of control when controlling for size of firm. In this section we provide comparative statics demonstrating how our model is consistent with these observations.

To do so, we first develop the notion of a skill premium within the continuous model. Similar to the skill premium defined for the discrete case in Section 3.1, we define the skill premium for a firm with X levels and span of control R to be the ratio of the average wage for managers to the wage for workers; that is,

$$SP(X, R) \equiv \frac{\int_{x>0} w(x, R) l(x, X, R) dx / \int_{x>0} l(x, X, R) dx}{1 - \lambda}, \quad (3)$$

where $l(x, X, R) = R^{X-x}$ is the number of employees at each level and $w(x, R)$ is the wage of a manager at level x working for a firm with span of control R . This aligns with the definition used in the data where skill premium was calculated as the ratio of average wages for managers to average wages of blue collar workers. Using this definition of skill premium, and the fact that

$$w(x, R) = MC_x = R^x e^{-x(1-\lambda)} (1 - \lambda),$$

we are able to prove that skill premium is increasing in the size of the span of control, holding constant the number of levels. We state this formally in Lemma 4 in Appendix A.

The intuition for this result is straightforward. The number of tasks passing through a manager in a given level is increasing in the number of direct reports. As the manager's span of control increases so does their marginal contribution to the firm.

A firm with X levels and span of control R will be of size

$$N(X, R) \equiv \int_{x>0} l(x, X, R) dx = \frac{R^X - 1}{\ln R}.$$

We can slightly abuse notation to calculate the relation between skill premium and the size of a firm. Consistent with what we observe in the data, we prove that this relation is positive (Lemma 5 in Appendix A).

This result comes directly from the fact that managers at higher levels or managers with a higher span of control have a higher marginal productivity than managers at lower levels or with lower R . A firm can get bigger either by adding more levels or by increasing the span of control, and either way the average productivity of managers increases. Since the productivity per worker is constant, the result is obtained.

At its core, our model is a model of human capital with an added layer of firm structure. Firm structure requires that firms hire specific ratios of employees across levels of human capital. Instead of being perfectly substitutable in production, imposing a structural hierarchy upon firms turns employees with varying levels of education into complements. Two workers cannot simply replace a manager but must work in conjunction with managers in order to accomplish additional tasks. Larger firms require more levels and ultimately a higher average level of human capital. Consequently, they pay a higher average skill premium.

4 General Equilibrium

We now embed the structure introduced in Section 3.2 within a static general equilibrium framework which provides one potential rationalization for our assumption on wages. Agents in our general equilibrium setup include consumers, firms and an education sector. Consumers supply labor inelastically and choose a quantity of education to accrue. The education a consumer receives determines the level of a firm the consumer is qualified to work at, the number of direct reports he is able to manage at a firm and the wage he earns. Consumers own the firms and use their labor income and profits from the firms to pay for education and the consumption good. Only by receiving education can an individual work as a manager. Firms operate a technology that converts completed tasks into the final good with heterogeneous productivities, and they choose an optimal number of levels of management and span of control. The education sector employs educators and is perfectly competitive.

4.1 Consumers

There is a measure L of ex-ante identical consumers. They are endowed with a unit of time and own an equal share of all firms. We will denote by Π the aggregate profits across all firms. Consumers choose whether to work in the education sector or the private sector by comparing the utility received from working in the education sector, V_E , with the utility

received from working in the private sector, V_P . We denote the utility function of consumers by $u(\cdot)$ and assume it is strictly increasing.

In order to obtain competence in educating others, educators require $h_E < 1$ units of instruction from other educators. Educators supply instruction in the education sector inelastically and receive a wage of w_E . The budget constraint for an educator is

$$c_E + w_E h_E = w_E + \frac{\Pi}{L}.$$

This implies that the utility an educator receives is

$$V_E \equiv u\left(w_E(1 - h_E) + \frac{\Pi}{L}\right).$$

Consumers in the private sector choose a level of education, characterized by a pair (x, R) . The quantity x denotes the level at which the consumer is able to manage at a firm. For example, an individual with a level of education $x = 1$ can work at the first level of management. R is the number of direct reports the consumer will receive at a firm. The amount of time required for an educator to train a consumer to work at level x with R direct reports is denoted by $h(x, R)$. Notice that there could be a trade-off between being a higher level, but getting fewer direct reports (high x and low R) and being at a lower level with more direct reports (low x and high R). Formally, the problem of a consumer in the private sector is:

$$\begin{aligned} V_P &\equiv \max_{x, R} u(c) \\ \text{s.t. } c + w_E h(x, R) &= w(x, R) + \frac{\Pi}{L} \end{aligned} \quad (4)$$

Since all consumers are ex-ante identical, it must be the case that consumers are indifferent between all choices in their optimal set. If we assume that there exists (x_E, R_E) such that the level of education required to work at level x_E and span of control R_E is equivalent to the level of education required to become an educator (i.e. $h(x_E, R_E) = h_E$), then the indifference requirement immediately implies that $w_E = w(x_E, R_E)$. This allows us to simplify the consumer's problem to be (4) since V_E is given by the utility obtained from choosing x_E .

Notice that in order to educate a manager to level x and span of control R we require $h(x, R)$ units of educator time. Further, each of those educators requires h_E units of educator time, and those educators also need h_E units of educator time, and so on. Therefore the

total time required in order to train a manager to level x and span of control R is

$$h(x, R) + h(x, R)h_E + (h(x, R)h_E)h_E + \cdots = h(x, R) \sum_{n=0}^{\infty} h_E^n = \frac{h(x, R)}{1 - h_E}.$$

We assume that

$$h(x, R) = (R^x e^{-x(1-\lambda)} - 1) (1 - h_E).$$

This assumption has the feature that the marginal cost of additional education is exactly equal to the marginal productivity in terms of tasks solved by a manager at level x for a firm with span of control R .¹¹

4.2 Firms

The homogeneous final good is produced by a continuum of firms. Each firm has an idiosyncratic technology parameter z distributed according to the density function $f(\cdot)$. Given z , the firm chooses a number of employee levels X in order to maximize profits. A firm with productivity z whose employees complete $\mathcal{T}$ total tasks produces the final good using technology $Y(\mathcal{T}, z)$. We make basic regularity assumptions on Y ; namely, $Y_{\mathcal{T}} > 0$, $Y_{\mathcal{T}\mathcal{T}} < 0$, $Y_z > 0$, and $Y(\mathcal{T}, z)$ satisfies Inada conditions on $\mathcal{T}$. In particular, this allows us to characterize a firm by its level of z .

The structure of the firm is rigid in the sense that a firm with X levels must employ R^{X-x} employees at each level $x \in [0, X]$ at the firm. In this sense, the firm faces a modified Leontief production function for which the proportions of employees hired at each level are fixed. Workers and managers are able to complete tasks as outlined in Section 3.2. Let $\mathcal{T}(X, R)$ denote the number of tasks completed by a firm with X levels and span of control R .¹² Profits for firm z are given by

$$\begin{aligned} \pi(z) &= \max_{X, R} Y(\mathcal{T}(X, R), z) - \int_x w(x, R) l(x, X, R) dx \\ \text{s.t. } l(x, X, R) &= R^{X-x}. \end{aligned} \tag{5}$$

Total profits across all firms are then $\Pi = \int_z \pi(z) f(z) dz$.

¹¹An equilibrium analysis under more general assumptions on $h(x, R)$ is provided in Appendix B.

¹²Recall that $\mathcal{T}(X, R) = R^X (1 - e^{-X(1-\lambda)})$.

4.3 Equilibrium

An equilibrium is a function which maps education levels and span of control to wages, $w : \mathcal{X} \times \mathcal{R} \rightarrow \mathbb{R}_+$, and a function which maps firm technology endowments to an optimal number of levels for the firm, $X^* : \mathcal{Z} \rightarrow \mathbb{R}_+$ and span of control, $R^* : \mathcal{Z} \rightarrow \mathbb{R}_+$, such that:

1. X^* and R^* solve the firm's problem (5);
2. consumers are indifferent between any education level which is employed with positive measure;
3. the total labor market clears

$$\int_z \int_x l(x, X^*(z), R^*(z)) \left(1 + \frac{h(x, R)}{1 - h_E}\right) f(z) dx dz = L;$$

4. the goods market clears

$$\int_z Y(\mathcal{T}(X^*(z), R^*(z)), z) f(z) dz = \int_{(x, R) \neq (x_E, R_E)} c(x, R) g(x \times R) d(x \times R) + M_E c(x_E, R_E),$$

where

$$\begin{aligned} M_E &\equiv \int_z \int_x l(x, X^*(z), R^*(z)) \frac{h(x, R^*(z))}{1 - h_E} f(z) dx dz \\ g(x \times R) &\equiv \frac{1}{L} \int_z l(x, X^*(z), R^*(z)) \mathbf{1}_{\{R^*(z)=R\}} f(z) dz \\ c(x, R) &= w(x, R) - w_E h(x, R) + \frac{\Pi}{L} \\ c(x_E, R_E) &= w_E (1 - h_E) + \frac{\Pi}{L}. \end{aligned}$$

4.4 Solving for wages

We denote by w_0 the wage rate for an unskilled worker on the bottom level of a firm, who has received no education. By equilibrium condition 2 in section 4.3 we know that consumers must be indifferent between obtaining any education level which is employed with positive measure. This implies that $w(x, R) = w_E h(x, R) + w_0$. It immediately follows that the wage for educators is $w_E = \frac{w_0}{1 - h_E}$. Therefore equilibrium object w_0 , which is determined through labor market clearing, together with the exogenous education function, $h(x, R)$, are sufficient for characterizing the entire wage schedule.¹³

¹³Specifically $w(x, R) = w_0 \left(\frac{h(x, R)}{1 - h_E} + 1 \right)$.

In Proposition 2 we demonstrate that a firm is indifferent between trading off height for width and that the firm's problem reduces to selecting a number of tasks to solve. In this very specific scenario, our model is isomorphic to a standard model with human capital. Due to the rigid nature of the firm's structure and the fact that $\ln R > 1 - \lambda$, the mapping from X to $\mathcal{T}$ is monotonic and we can simplify the firm's problem to choosing the optimal number of tasks to be solved. In Proposition 2 in Appendix A we formally prove that the firm's problem collapses to

$$\max_{\mathcal{T}} Y(\mathcal{T}, z) - w_0 \mathcal{T}. \quad (6)$$

This immediately implies that a firm with a higher technology parameter z will solve more tasks. Since tasks are monotonically increasing in the number of levels, it is also straightforward that the number of levels in a firm is increasing in z . The continuous case offers an analytical advantage over the discrete case since solutions do not depend on cutoff values for z . The continuity in selection of tasks along z allows us to demonstrate the existence of the equilibrium which we formalize with Proposition 3 in Appendix A.

In particular, if $Y(\mathcal{T}, z) = z\mathcal{T}^\alpha$, then

$$\mathcal{T}^*(z) = \left(\frac{\alpha z}{w_0} \right)^{\frac{1}{1-\alpha}}$$

for each z and

$$w_0 = \alpha \left(\frac{\mathbf{E} \left[z^{\frac{1}{1-\alpha}} \right]}{L(1-\lambda)} \right)^{1-\alpha}$$

clears the labor market.

That marginal productivity is equal to marginal cost at an employee level is not a requirement for our model. Due to the ex-ante homogeneity of consumers, the supply of managers at any level of sophistication is perfectly elastic at the rate which exactly compensates for their cost of education. The costs of education will directly influence a firm's choice in regards to the optimal ratio of width and height for a firm with specific technology z . While this is an interesting implication of our model, it is not the focus of the paper and we relegate further discussion to Appendix B.

5 Conclusion

In this paper we use recent data from the Chilean Manufacturing Survey to document that, consistent with previous findings, skill premium is positively related with the size of firms.

We exploit the rich information we have on different types of skilled workers to estimate the relationship between skill premia by worker type and firm size. We find that the effect of size on skill premium is much greater for managers and owners than for other categories of skilled workers. From this we hypothesize that the organizational structure of the firm is important in explaining the positive relationship between size and skill premium. Further findings on the positive correlation between the ratio of workers to managers and skill premium suggest that span of control plays a key role in a firm's organizational structure.

We build on the discrete structural models in the literature (Garicano, 2000; Caliendo and Rossi-Hansberg, 2012) by developing a simple continuous version of the model which maps higher levels of management to higher levels of human capital. The model allows for a closed form analytical solution and provides comparative statics which make specific predictions about the effects of organizational structure on its wage structure; namely, the two facts in the data outlined above.

The current literature on skill biased technical change is often linked with recent advances in computing technology. However, exactly how computing technology disproportionately augments the wages of the skilled is widely debated. Our data and model offer one potential avenue for further exploration, to the extent computing technology extends the span of control of managers and owners, computing technology will be correlated with augmented wages for skilled workers.

Appendix A Mathematical appendix

Lemma 1. *Consider a firm with X levels and span of control R . Each manager in level $x \in \{1, \dots, X\}$ solves a measure $(R\lambda)^x(1 - \lambda)$ of tasks. The total measure of tasks completed by the firm is $R^X(1 - \lambda^{X+1})$.*

Proof. We will first prove by induction on x that each manager at level $x > 0$ is able to solve $(R\lambda)^x(1 - \lambda)$ tasks. Consider first the measure of tasks that workers solve: By assumption each worker receives a measure one of tasks of which he is unable to solve a fraction λ which he passes on to managers in level $x = 1$. Each manager at level $x = 1$ receives R direct reports, so they receive a measure $R\lambda$ of tasks, out of which they solve a fraction $1 - \lambda$ and pass on the remaining to managers at level $x = 2$. Hence managers at level $x = 1$ solve $R\lambda(1 - \lambda)$ tasks.

Consider a manager at level x and assume that he solves a measure $(R\lambda)^x(1 - \lambda)$ of tasks. A manager at level $x + 1$ receives R direct reports, each of which passes on a measure $\lambda(R\lambda)^x$ of unsolved tasks. Of this measure she is able to solve a fraction $1 - \lambda$. This implies that the manager at level $x + 1$ solves $(R\lambda)^{x+1}(1 - \lambda)$ tasks.

We now prove that the total measure of tasks completed by a firm is $R^X(1 - \lambda^{X+1})$. Since

there are R^{X-x} managers at level x , the total number of tasks $\mathcal{T}_X$ completed by the firm is

$$\begin{aligned}\mathcal{T}_X &= \sum_{x=0}^X R^{X-x} (R\lambda)^x (1-\lambda) \\ &= R^X (1 - \lambda^{X+1}).\end{aligned}$$

□

Lemma 2. *If $R\lambda > 1$ then $SP_{X+1,R} > SP_{X,R}$ and $SP_{X,R+1} > SP_{X,R}$.*

Proof. First we will prove that $SP_{X+1,R} > SP_{X,R}$. For this it is sufficient to prove that

$$\frac{R^{X+1}}{R^{X+1} - 1} (1 - \lambda^{X+1}) > \frac{R^X}{R^X - 1} (1 - \lambda^X),$$

which is equivalent to proving that

$$R \frac{\sum_{x=0}^X \lambda^x}{\sum_{x=0}^{X-1} \lambda^x} > \frac{\sum_{x=0}^X R^x}{\sum_{x=0}^{X-1} R^x}. \quad (\text{A1})$$

(A1) holds if and only if

$$\begin{aligned}R \left(\sum_{x=0}^{X-1} R^x \right) \left(\sum_{x=0}^X \lambda^x \right) &> \left(\sum_{x=0}^X R^x \right) \left(\sum_{x=0}^{X-1} \lambda^x \right) \\ \iff \lambda^X \sum_{x=1}^{X-1} R^x &> \sum_{x=0}^{X-1} \lambda^x \\ \iff \sum_{x=0}^{X-1} \lambda^x (R\lambda)^{X-x} &> \sum_{x=0}^{X-1} \lambda^x.\end{aligned} \quad (\text{A2})$$

Since $R\lambda > 1$, λ^x on the left hand side of (A2) is multiplied by a term greater than one for all $x \in \{0, \dots, X-1\}$. Therefore the left hand side is strictly greater than the right hand side and the result follows.

Now we prove that $SP_{X,R+1} > SP_{X,R}$. For this it is sufficient to prove that

$$R \frac{(R+1)^X}{(R+1)^X - 1} > (R-1) \frac{R^X}{R^X - 1},$$

which is equivalent to proving that

$$R(R^X - 1)(R+1)^X > (R-1)R^X((R+1)^X - 1). \quad (\text{A3})$$

(A3) holds if and only if

$$(R^X - R)(R+1)^X + (R^{X+1} - R^X) > 0,$$

which holds if $R > 1$. Notice that $R\lambda > 1$ and $\lambda \in (0, 1)$ imply $R > 1$, which completes the proof. $\square$

Lemma 3. *There is a continuum of measure $\frac{1}{\ln R}(R^x - 1)$ of employees in area of command $AC(x, n)$. If there is a measure 1 of CEOs, then there is a measure $\frac{R^X - 1}{\ln R}$ of managers in a firm with X levels and span of control R .*

Proof. The measure of managers in $AC(x, n)$ is given by

$$\int_{(x', n') \in AC(x, n)} dx' dn' = \frac{1}{\ln R}(R^x - 1)$$

$\square$

Proposition 1. *Suppose each worker receives a measure one of tasks. Let $\mathcal{T}(x, R)$ be the number of tasks that are solved in the area of command $AC(x)$. Then*

$$\mathcal{T}(x, R) = R^x (1 - e^{-x(1-\lambda)}).$$

Assume that there is a measure 1 of CEOs. Then the total number of tasks solved by a firm with X levels and R reports is $R^X(1 - e^{-X(1-\lambda)})$.

Proof. We will prove the result by first considering levels of width Δ and calculating by induction the total number of tasks solved in $AC(x)$ as $\Delta \rightarrow 0$. In each level $x' \in [0, x]$ of area of command $AC(x)$ there are $R^{x-x'}$ employees. Workers in $AC(x)$ receive a measure 1 of tasks and are able to solve a fraction $\Delta(1 - \lambda)$ of them. Each manager in level Δ receives $R^\Delta(1 - \Delta(1 - \lambda))$ tasks and is able to solve a fraction $\Delta(1 - \lambda)$. Since there is a measure $R^{x-\Delta}$ of managers at level Δ , the total measure of tasks that managers at level Δ solve is $R^x(1 - \Delta(1 - \lambda))\Delta(1 - \lambda)$. Now consider managers at level 2Δ . Each employee in level 2Δ manages a width of R^Δ direct reports, each of whom passes on $R^\Delta(1 - \Delta(1 - \lambda))^2$ unsolved tasks. Each manager in level 2Δ is able to solve a fraction $\Delta(1 - \lambda)$ of those tasks. There is a measure $R^{x-2\Delta}$ of managers at level 2Δ , so the total measure of tasks that these managers solve is $R^x(1 - \Delta(1 - \lambda))^2\Delta(1 - \lambda)$.

Now assume that managers at level $i\Delta$ solve a measure $R^{i\Delta}(1 - \Delta(1 - \lambda))^i\Delta(1 - \lambda)$ of tasks. There are $R^{x-i\Delta}$ managers at this level. So the total measure of tasks solved by managers at this level is $R^x(1 - \Delta(1 - \lambda))^i\Delta(1 - \lambda)$. Each employee at level $(i+1)\Delta$ manages a width of R^Δ direct reports, each of whom passes on $R^{i\Delta}(1 - \Delta(1 - \lambda))^{i+1}$ unsolved tasks. Each manager in level $(i+1)\Delta$ is able to solve a fraction $\Delta(1 - \lambda)$ of those tasks. Since there are $R^{x-(i+1)\Delta}$ managers at level $i+1$, the total measure of tasks solved by these managers is $R^x(1 - \Delta(1 - \lambda))^{i+1}\Delta(1 - \lambda)$. Therefore the total measure of tasks solved in $AC(x)$ is

$$\begin{aligned} \mathcal{T}(x, R) &\equiv \lim_{\Delta \rightarrow 0} (1 - \lambda) R^x \Delta \sum_{i=0}^{\lfloor \frac{x}{\Delta} \rfloor} (1 - \Delta(1 - \lambda))^i \\ &= \lim_{\Delta \rightarrow 0} R^x (1 - (1 - \Delta(1 - \lambda))^{\lfloor \frac{x}{\Delta} \rfloor + 1}) \\ &= R^x (1 - e^{-x(1-\lambda)}), \end{aligned}$$

where the last line is due to the fact that

$$\lim_{\Delta \rightarrow 0} \frac{x}{\Delta} \ln(1 - \Delta(1 - \lambda)) = -x(1 - \lambda).$$

Finally, a firm with X levels and R reports is $AC(X)$. Proposition (1) implies that $\mathcal{T}(X) = R^X(1 - e^{-X(1-\lambda)})$. $\square$

Lemma 4. $\frac{\partial SP}{\partial R} > 0$.

Proof. Since

$$\begin{aligned} w(x, R) &= MC_x = R^x e^{-x(1-\lambda)}(1 - \lambda) \\ l(x, R, X) &= R^{X-x}, \end{aligned}$$

we can write

$$SP(X, R) = \ln R \frac{1}{1 - \lambda} \frac{R^X}{R^X - 1} (1 - e^{-X(1-\lambda)}).$$

Let $g(R) = \ln R \frac{R^X}{R^X - 1}$. It is sufficient to prove that $g'(R) > 0$, where

$$g'(R) = \frac{R^{X-1}}{(R^X - 1)^2} (R^X - 1 - \ln R^X).$$

The term $\frac{R^{X-1}}{(R^X - 1)^2} > 0$ since $R > 0$. Since $y - 1 > \ln y$ for any $y > 0$. By letting $y = R^X$, the result follows. $\square$

Lemma 5. *Skill premium is increasing in the size of a firm:* $\frac{\partial SP}{\partial N} > 0$.

Proof. From the proof of Lemma 4 we have

$$SP(X, R) = \ln R \frac{1}{1 - \lambda} \frac{R^X}{R^X - 1} (1 - e^{-X(1-\lambda)}).$$

To see that $\frac{\partial SP}{\partial N} > 0$, first realize that $\frac{\partial SP}{\partial N} = \frac{\partial SP}{\partial X} \frac{\partial X}{\partial N}$. The number of workers at a firm is $N = \frac{R^X - 1}{\ln R}$. Solving for X we get $X = \frac{\ln(N \ln R + 1)}{\ln R}$. From this we can calculate $\frac{\partial X}{\partial N} = \frac{1}{N \ln R + 1} > 0$.

Showing that $\frac{\partial SP}{\partial X} > 0$ is equivalent to showing that $g'(X) > 0$ where

$$g(X) \equiv \frac{R^X}{R^X - 1} (1 - e^{-X(1-\lambda)}).$$

Since

$$g'(X) = \frac{R^X}{e^{(1-\lambda)X} (R^X - 1)^2} ((1 - \lambda)(R^X - 1) - \ln R (e^{X(1-\lambda)} - 1)),$$

it is sufficient to prove that $\hat{g}(X) \equiv (1 - \lambda)(R^X - 1) - \ln R (e^{X(1-\lambda)} - 1) > 0$ for all $X > 0$. Notice that $\hat{g}(0) = 0$. Also,

$$\hat{g}'(X) = (1 - \lambda) \ln R (R^X - e^{X(1-\lambda)}),$$

$\square$

Proposition 2. *If $\ln R > 1 - \lambda$ the firm's problem becomes:*

$$\max_{\mathcal{T}} Y(\mathcal{T}, z) - w_0 \mathcal{T}.$$

Proof. Recall from Proposition 1 that the number of tasks a firm solves is given by $\mathcal{T}(X, R) = R^X (1 - e^{-X(1-\lambda)})$. Equilibrium condition 2 requires that consumers be indifferent between obtaining any education levels which are employed with positive measure. This implies that $w(x, R) = w_E h(x, R) + w_0$. Now, for an educator it must hold that $w_E = w_E h_E + w_0$, so $w_E = \frac{w_0}{1-h_E}$. Therefore

$$w(x) = \frac{w_0}{1-h_E} h(x, R) + w_0. \quad (\text{A4})$$

Plugging (A4) and the labor requirements at each level into the firm's problem we get:

$$\begin{aligned} & \max_X Y(\mathcal{T}(X), z) - \int_x R^{X-x} \left(\frac{w_0}{1-h_E} h(x, R) + w_0 \right) dx \\ &= \max_X Y(\mathcal{T}(X), z) - R^X \int_x w_0 e^{-x(1-\lambda)} dx \\ &= \max_X Y(\mathcal{T}(X), z) - w_0 \mathcal{T}(X). \end{aligned} \quad (\text{A5})$$

$\ln R > 1 - \lambda$ implies that $\mathcal{T}(X, R)$ is strictly increasing in X .¹⁴ Therefore (A5) is equivalent to

$$\max_{\mathcal{T}} Y(\mathcal{T}, z) - w_0 \mathcal{T}.$$

□

Proposition 3. *In this economy there exists an equilibrium.*

Proof. Proposition 2 implies that the solution to (6) is sufficient to derive a function $X^* : \mathcal{Z} \rightarrow \mathbb{R}_+$ that satisfies the equilibrium conditions given wages.

Denote by $\mathcal{T}^*(z, w_0)$ the solution to (6). The regularity conditions on $Y(\mathcal{T}, z)$ imply that $\mathcal{T}^*(z, w_0)$ is strictly decreasing in w_0 , $\lim_{w_0 \rightarrow \infty} \mathcal{T}^*(z, w_0) = 0$ and $\lim_{w_0 \rightarrow 0} \mathcal{T}^*(z, w_0) = \infty$ since $\mathcal{T}^*(z, w_0)$ satisfies

$$Y_{\mathcal{T}}(\mathcal{T}^*(z, w_0), z) = w_0.$$

Denote by $X^*(z, w_0)$ the value of X that achieves $\mathcal{T}^*(z, w_0)$. Notice that $X^*(z, w_0)$ is also strictly decreasing in w_0 . Let

$$\begin{aligned} H(w_0) &\equiv \int_z \int_x l(x, X^*(z, w_0)) \left(1 + \frac{h(x)}{1-h_E} \right) f(z) dx dz \\ &= \int_z \int_x R^{X^*(z, w_0)-x} R^x e^{-x(1-\lambda)} f(z) dx dz \\ &= \frac{1}{1-\lambda} \int_z \mathcal{T}^*(z, w_0) f(z) dz. \end{aligned}$$

¹⁴ $\mathcal{T}'(X) = R^X (\ln R - e^{-X(1-\lambda)} (\ln R - (1-\lambda)))$. Since $\ln R > 1 - \lambda$, $1 - \lambda > 0$ and $e^{-X(1-\lambda)} < 1$, then $\mathcal{T}'(X) > 0$.

Then the regularity conditions on $Y(\mathcal{T}, w_0)$ imply that $H(w_0)$ is strictly decreasing in w_0 , $\lim_{w_0 \rightarrow \infty} H(w_0) = 0$ and $\lim_{w_0 \rightarrow 0} H(w_0) = \infty$. Since the labor market clearing condition implies $L = H(w_0)$, the result follows. $\square$

Appendix B General h

In this section we set up and characterize the problem each firm faces without imposing any assumption on $h(x, R)$, other than it being strictly increasing in both of its arguments. Recall the problem that the firm solves, (5):

$$\pi(z) = \max_{X, R} Y(\mathcal{T}(X, R), z) - w_0 \int_x R^{X-x} \left(1 + \frac{h(x, R)}{1 - h_E} \right) dx,$$

where we are using the fact that in equilibrium, consumers are indifferent between any education level which is employed with positive measure, and, thus, $w(x, R) = w_E h(x, R) + w_0$, that $w_E = \frac{w_0}{1 - h_E}$ and that the number of workers in a given level is equal to R^{X-x} . To simplify notation, we will refer to the costs the firm faces as TC :

$$TC \equiv w_0 \int_x R^{X-x} \left(1 + \frac{h(x, R)}{1 - h_E} \right) dx.$$

To find a solution for this problem, the firm will choose R and X such that on the margin the firm is indifferent between increasing R or increasing X . That is, the marginal cost per task of increasing X and R must satisfy

$$\frac{\mathcal{T}_X}{TC_X} = \frac{\mathcal{T}_R}{TC_R} = \frac{1}{TC_{\mathcal{T}}}. \quad (\text{B6})$$

Additionally, the marginal benefit per task must equal the marginal cost per task:

$$Y_{\mathcal{T}} = TC_{\mathcal{T}}.$$

Since the choice of $h(X, R)$ we use in the body of the paper causes total costs to equal total tasks times w_0 ,¹⁵ then condition (B6) is satisfied trivially.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis**, “High Wage Workers and High Wage Firms,” *Econometrica*, 1999, 67 (2), pp. 251–333.
- Beggs, A. W.**, “Queues and Hierarchies,” *Review of Economic Studies*, 2001, 68 (2), 297–322.

¹⁵See Proposition 2 in Appendix A.

- Bernard, Andrew B. and J. Bradford Jensen**, “Exceptional exporter performance: cause, effect, or both?,” *Journal of International Economics*, February 1999, 47 (1), 1–25.
- Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen**, “The Distinct Effects of Information Technology and Communication Technology on Firm Organization,” *Management Science*, December 2014, 60 (12), 2859–2885.
- Brown, Charles and James Medoff**, “The Employer Size-Wage Effect,” *Journal of Political Economy*, October 1989, 97 (5), 1027–59.
- Bustos, Paula**, “Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms,” *American Economic Review*, February 2011, 101 (1), 304–40.
- Caliendo, Lorenzo and Esteban Rossi-Hansberg**, “The Impact of Trade on Organization and Productivity,” *The Quarterly Journal of Economics*, 2012, 127 (3), 1393–1467.
- , **Ferdinando Monte, and Esteban Rossi-Hansberg**, “The Anatomy of French Production Hierarchies,” *Journal of Political Economy*, 2015, 123 (4), 809 – 852.
- Davis, Steve J., John Haltiwanger, Lawrence F. Katz, and Robert Topel**, “Wage Dispersion between and within U.S. Manufacturing Plants, 1963–86,” *Brookings Papers on Economic Activity. Microeconomics*, 1991, 1991, pp. 115–200.
- Dunne, Timothy and Jr Schmitz James A**, “Wages, Employment Structure and Employer Size-Wage Premia: Their Relationship to Advanced-Technology Usage at US Manufacturing Establishments,” *Economica*, February 1995, 62 (245), 89–107.
- , **Lucia Foster, John C. Haltiwanger, and Kenneth Troske**, “Wage and Productivity Dispersion in U.S. Manufacturing: The Role of Computer Investment,” IZA Discussion Papers 563, Institute for the Study of Labor (IZA) August 2002.
- Fox, Jeremy T.**, “Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching,” *Journal of Labor Economics*, 01 2009, 27 (1), 83–126.
- Garicano, Luis**, “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, October 2000, 108 (5), 874–904.
- **and Esteban Rossi-Hansberg**, “Organization and Inequality in a Knowledge Economy,” *The Quarterly Journal of Economics*, 2006, 121 (4), 1383–1435.
- **and Thomas N. Hubbard**, “Managerial Leverage Is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyers’ Human Capital,” *Journal of Law and Economics*, February 2007, 50 (1), 1–43.
- **and –**, “Specialization, Firms, and Markets: The Division of Labor within and between Law Firms,” *Journal of Law, Economics and Organization*, October 2009, 25 (2), 339–371.
- Grobovšek, Jan**, “Management and Aggregate Productivity,” 2013. Mimeo, University of Edinburgh.

- Haskel, Jonathan**, “The Skilled Wage Premium and the Growth of Small Firms in UK Manufacturing,” *LABOUR*, 07 1998, *12* (2), 221–238.
- Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding**, “Inequality and Unemployment in a Global Economy,” *Econometrica*, 07 2010, *78* (4), 1239–1283.
- Idson, Todd L. and Walter Y. Oi**, “Workers Are More Productive in Large Firms,” *American Economic Review*, May 1999, *89* (2), 104–108.
- Katz, Lawrence F. and Lawrence H. Summers**, “Can Interindustry Wage Differentials Justify Strategic Trade Policy?,” in “Trade Policies for International Competitiveness” NBER Chapters, National Bureau of Economic Research, Inc, October 1989, pp. 85–124.
- Maio, Michael**, “Essays in International Trade and Development.” PhD dissertation, University of Minnesota 2014.
- Nikolowa, Radoslaw**, “Supply of skilled labour and organizational change,” *Labour Economics*, June 2010, *17* (3), 514–522.
- Pavcnik, Nina**, “What explains skill upgrading in less developed countries?,” *Journal of Development Economics*, 2003, *71* (2), 311–328.
- Yeaple, Stephen Ross**, “A simple model of firm heterogeneity, international trade, and wages,” *Journal of International Economics*, January 2005, *65* (1), 1–20.
- Zweimüller, Josef and Rudolf Winter-Ebmer**, “On-the-job-training, job search and job mobility,” *Swiss Journal of Economics and Statistics (SJES)*, December 2003, *139* (IV), 563–576.