Pontifícia Universidade Católica do Rio de Janeiro

Departamento de Economia

Monografia de Final de Curso

Wage returns to skills in Brazil

Igor Rigolon La-Cava Veiga

1910827

Orientador: Gustavo Gonzaga

Rio de Janeiro,

Dezembro de 2022

Pontifícia Universidade Católica
do Rio de Janeiro

Igor Rigolon La-Cava Veiga

Wage returns to skills in Brazil

Monografia de Final de Curso

Orientador: Gustavo Gonzaga

Declaro que o presente trabalho é de minha autoria e que não recorri, para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

Rio de Janeiro,

Dezembro de 2022

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

Abstract

Using the newly developed Brazilian Qualifications Framework, I measure the levels of routine, cognitive, physical, and practical skills that each occupation requires from workers. I estimate the returns to these skills in the formal labor market, and find that workers are paid less in occupations more intensive in routine tasks. Workers in more cognitive-intensive occupations earn considerable wage premia. This result holds for workers within the same occupation, where those with more previous experience in cognitive skills are paid more. Workers gain higher wages when switching to an occupation more intense in cognitive skills, but lose wages if the new occupation has a skill profile too distant from the previous one.

Keywords: labor economics, skills, tasks

Contents

Contents		5
	List of Figures	6
	List of Tables	7
1	AN OVERVIEW OF SKILLS AND TASKS	9
1.1	Introduction	9
1.2	Skills and Job Polarization	10
1.3	Tasks	15
2	THE SKILL PROFILE OF BRAZILIAN OCCUPATIONS	17
2.1	Data Sources	17
2.2	Measuring routine tasks	19
2.3	Principal Component Analysis	21
2.3.1	Previous uses in the literature	22
2.3.2	Descriptive statistics for the skill indices	23
3	EMPIRICAL STRATEGY	29
3.1	Returns to skills	29
3.2	Skill experience	30
3.3	Occupational distance and wages	31
4	RESULTS	33
4.1	Returns to skills	33
4.2	Skill experience	34
4.3	Occupational distance and wages	35
5	CONCLUSION	39
	Bibliography	41
A	APPENDIX	43
A.1	Autonomy	43
A.2	Abstract Reasoning	44
A.3	Interaction	45
A.4	Technology	46

List of Figures

Figure 1 – Education by qualification level . . . 11
Figure 2 – Distribution of qualification levels . . . 12
Figure 3 – Job polarization . . . 13
Figure 4 – Qualification levels by ISCO occupation group . . . 14
Figure 5 – Coverage of the QBQ . . . 18
Figure 6 – Routine work by qualification level . . . 19
Figure 7 – Routine work by education level . . . 20
Figure 8 – Routine work by ISCO occupation group . . . 20
Figure 9 – Changes in mean routine levels . . . 21
Figure 10 – PCA loadings of each skill index . . . 26
Figure 11 – Skill levels by education . . . 27
Figure 12 – Skill levels by ISCO occupation group . . . 27
Figure 13 – Skill levels by wage . . . 28
Figure 14 – Change in skill levels . . . 28
Figure 15 – Autonomy by qualification level . . . 43
Figure 16 – Autonomy by ISCO occupational group . . . 44
Figure 17 – Abstract reasoning by qualification level . . . 45
Figure 18 – Abstract reasoning by ISCO occupational group . . . 45
Figure 19 – Interaction by ISCO occupational group . . . 46
Figure 20 – Technology by ISCO occupational group . . . 46

List of Tables

Table 1 – Correlation matrix between skill indices 25

Table 2 – Wage returns to skills 33

Table 3 – Wage returns to skills by education level 34

Table 4 – Wage returns to skill experience 35

Table 5 – Wage returns to job switches 36

Table 6 – Wage returns to job switches by education level 37

1 An overview of skills and tasks

1.1 Introduction

The traditional Mincerian explanatory variables for wages are education and experience, both of which attempt to measure how much general human capital a worker has accumulated. Given a level of general human capital, some workers are still more at risk of losing their jobs depending on the particular type of work they do (Autor and Dorn 2013). This calls for an attention towards forms of specific human capital that workers build up, and which only apply to their current place in the labor market.

Human capital can be sector-specific, in the form of knowledge obtained from experience within a particular sector; it can be firm-specific, allowing the worker to best maneuver a certain corporate structure; or it can be the ability to perform certain types of tasks, which I will focus on. Poletaev and Robinson (2008) find that the skill profile of an occupation best encapsulates both the levels of sector- and firm-specific human capital, accounting for most of the wage changes upon an occupation switch.

I use the recent Brazilian Qualifications Framework, *Quadro Brasileiro de Qualificações* (QBQ), to construct aggregate skill indices, measuring the intensity of routine, cognitive, physical, and practical skills required to work in each occupation. I estimate the returns to each of these skill measures in the Brazilian formal labor market. I also compute each worker's cumulative experience in each skill, and find that differences in skill background are relevant to explain differences in wages, even for workers in the same occupation. Finally, I estimate how much moving to an occupation with a vastly different skill set impacts wages. My main estimates show that workers performing more routine tasks earn lower wages. Those relying more heavily on cognitive skills earn a sizeable wage premium, which is larger for higher education levels.

According to Acemoglu and Autor (2011), skills are a worker's unobservable endowments that guide how well they can perform each activity. Tasks are the activities they actually execute while working. Depending on the occupation a worker finds themself in, they may perform different sets of tasks regardless of their inherent skills or preferences. As a result, tasks come from an interaction between a worker's set of skill endowments, and how much the market rewards each skill (Autor 2013).

Each occupation is treated as a bundle of tasks performed by their workers, approximately observed from the QBQ skill requirements. The main pitfall of this approach is the static nature of the data, which sees occupations as unchanging and uniform. This is a feature not exclusive to the QBQ, but is also common to the most widely used databases for measuring tasks in the US. Both the large official records of the Dictionary of Occupational Titles (DOT) and the O*NET consist of still pictures of occupations, updated only occasionally.

Using survey data, Autor and Handel (2013) find high variance in tasks performed by workers in the same occupations, suggesting that tasks have an important individual-level component, not solely bound by the occupation. In a similar vein, Atalay et al. (2020) led a shift in the literature by employing Natural Language Processing to extract task measures from job ads over decades. Wielding long-term ever-changing data, in which tasks vary both within occupations and over time, they found that a large portion of the change in tasks performed happened due to changes in the occupations themselves, which were not captured by most data sources. Besides being heterogeneous between workers in the same occupation, the task content of each occupation also changes significantly over time.

To address the problem of within-occupation task heterogeneity, I adopt a measure of each individual's accumulated experience in a particular skill used by Gonzaga and Guanziroli (2019). Since workers in the same occupation can have vastly different backgrounds, this serves as a closer approximation of each individual's skill set than the skills used in their current occupation. I find that the returns to skill experience are similar to the returns to skills used in the current occupation, even when controlling for the occupation itself. Workers previously in more cognitive-intensive occupation earn a considerable wage premium, even compared to workers in the same current occupation.

When a worker switches occupations, Poletaev and Robinson (2008) explain that the effect on wages is twofold: they earn higher wages by moving to an occupation that requires skills more highly valued in the labor market, but receive a wage penalty for losing specific human capital, which aided them in the previous occupation and is now of less use. I calculate the distance between the skill profiles of all pairs of occupations, and find that switches further away in the skill spectrum are linked to a decrease in wages, even when controlling for current skills.

1.2 Skills and Job Polarization

In this section, I highlight the empirical hurdles that led researchers into understanding occupations through the lens of tasks. The preceding paradigm of the canonical model described by Acemoglu and Autor (2011) compared workers of differing education levels to explain trends in employment and wage premia. Under technical change, which was perceived as skill-biased, the model predicted a substitution of low-skilled labor by machines, whereas high-skilled labor was made more valuable by the higher technological efficiency.

This turned out not to be consistent across low-skill occupations, as Autor and Dorn (2013) show that certain types of low-skill, yet hard to replace occupations faced rapid growth in the service sector. Meanwhile, the occupations most heavily replaced by technology were comprised of middle-skilled clerical work. The conjunction of increased demand for high-skilled workers, along with a decline of the middle-skilled labor share, but a countervailing rise in low-skilled occupations constituted the phenomenon named job polarization.

To answer whether these findings have also manifested themselves in Brazil, I employ the QBQ's scale of qualification levels ranging discretely from 1 to 5. The structure of this database is explained in more detail in section 2.1. Figure 1 shows the composition of each of these qualification levels in regards to their educational attainment. The main takeaway is that this measure of skill level, based on the activities and requirements of each occupation, is tightly linked to higher education levels. Education levels have increased across the board in this time period. The panel on the right, referring to 2019, indicates that the top two qualification levels are the ones with a significant proportion of college-educated workers, while the bottom three are mainly composed of high school graduates, with the first skill level having a relevant portion of workers without a high school degree.

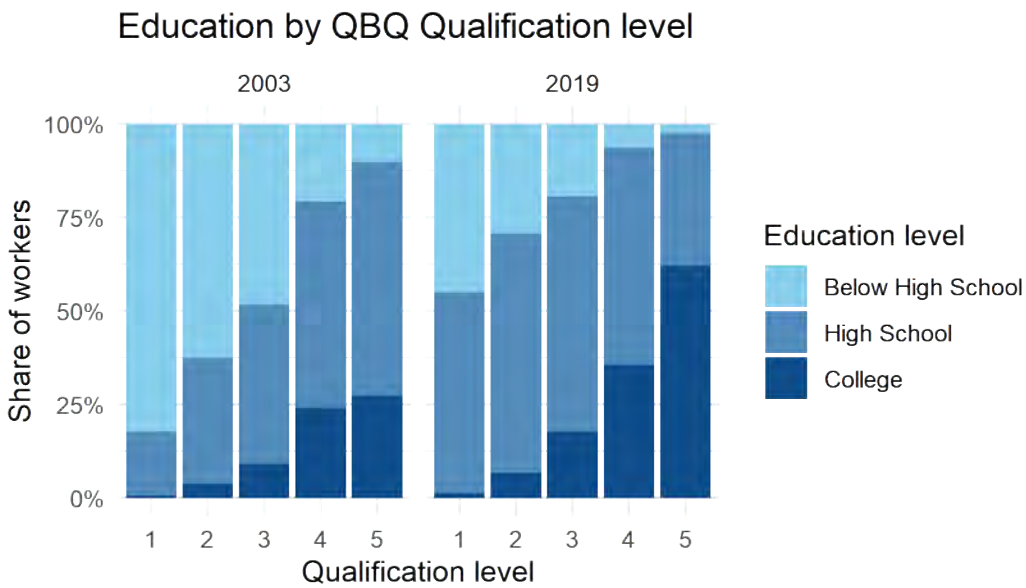


Figure 1

As an initial picture of the overall employment changes in Brazil, Figure 2 shows which proportion of workers fit into each of the QBQ's qualification levels.

It in fact illustrates a timid but consistent pattern of job polarization, where employment rose both at the top two qualification levels and, most notably, at the very lowest one. Meanwhile, occupations of qualification levels two and three, towards the middle of the pack, declined.

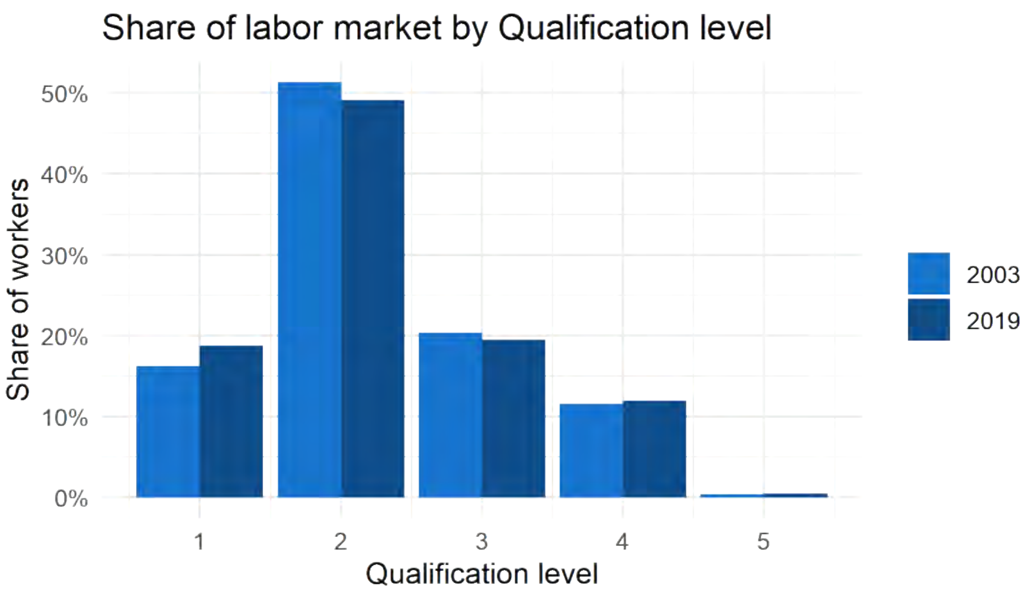


Figure 2

In more canonical presentation, following Autor and Dorn (2013), Figure 3 also plots changes of employment share, but with the x-axis now ranking occupations by their 2003 wage percentiles. This is taken as a proxy for skill, which, besides double checking what was seen for qualification levels, has the perk of being in a continuous scale. The curve follows a U-shape characteristic of job polarization, but lacks growth towards the bottom of the skill spectrum. It convincingly points to a relative decline in middle-skilled occupations.

To disaggregate the changes within each qualification level, I employ a mapping of the Brazilian Occupation Code from the *Classificação Brasileira de Ocupações* (CBO) into the ISCO-88, as originally done by Muendler et al. (2004) and adapted to the current edition of the CBO, lauched in 2002, by Corseuil, Foguel, and Gonzaga (2019). They then use the ISCO codes to map each occupation into 7 occupational categories, which I use to break down each qualification level in Figure 4, with the omission of Agriculture. As the QBQ is yet to classify occupations at the highest skill levels, there are also no Managerial occupations in the sample.

The changes at the lowest skill level mostly arise from a balance between growing service and sales occupations, along with declining production and clerical jobs, which lines up with the findings of Autor and Dorn (2013). This remains true for skill level two, but with much fainter growth in service occupation, which

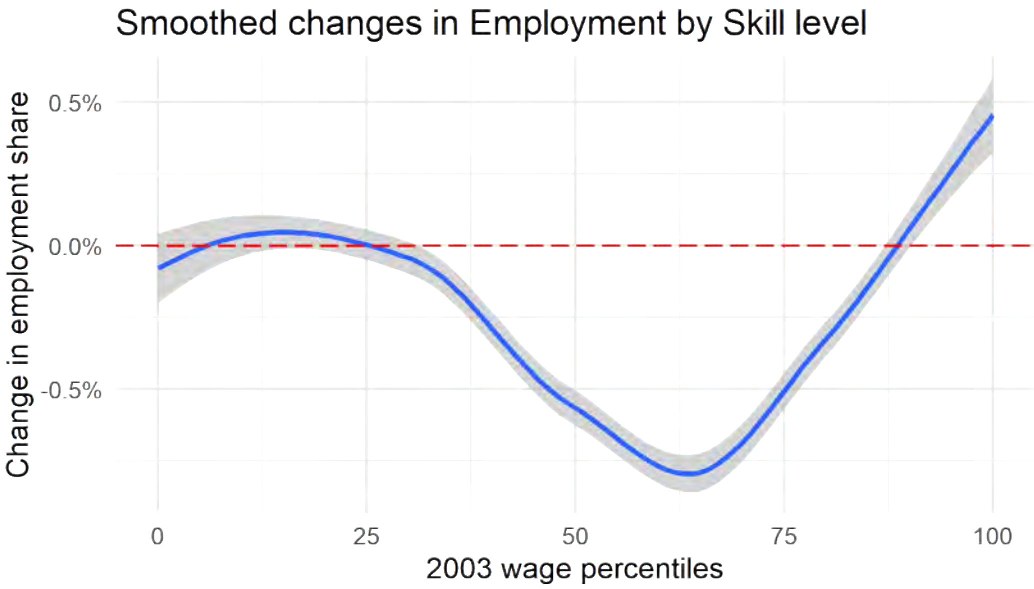
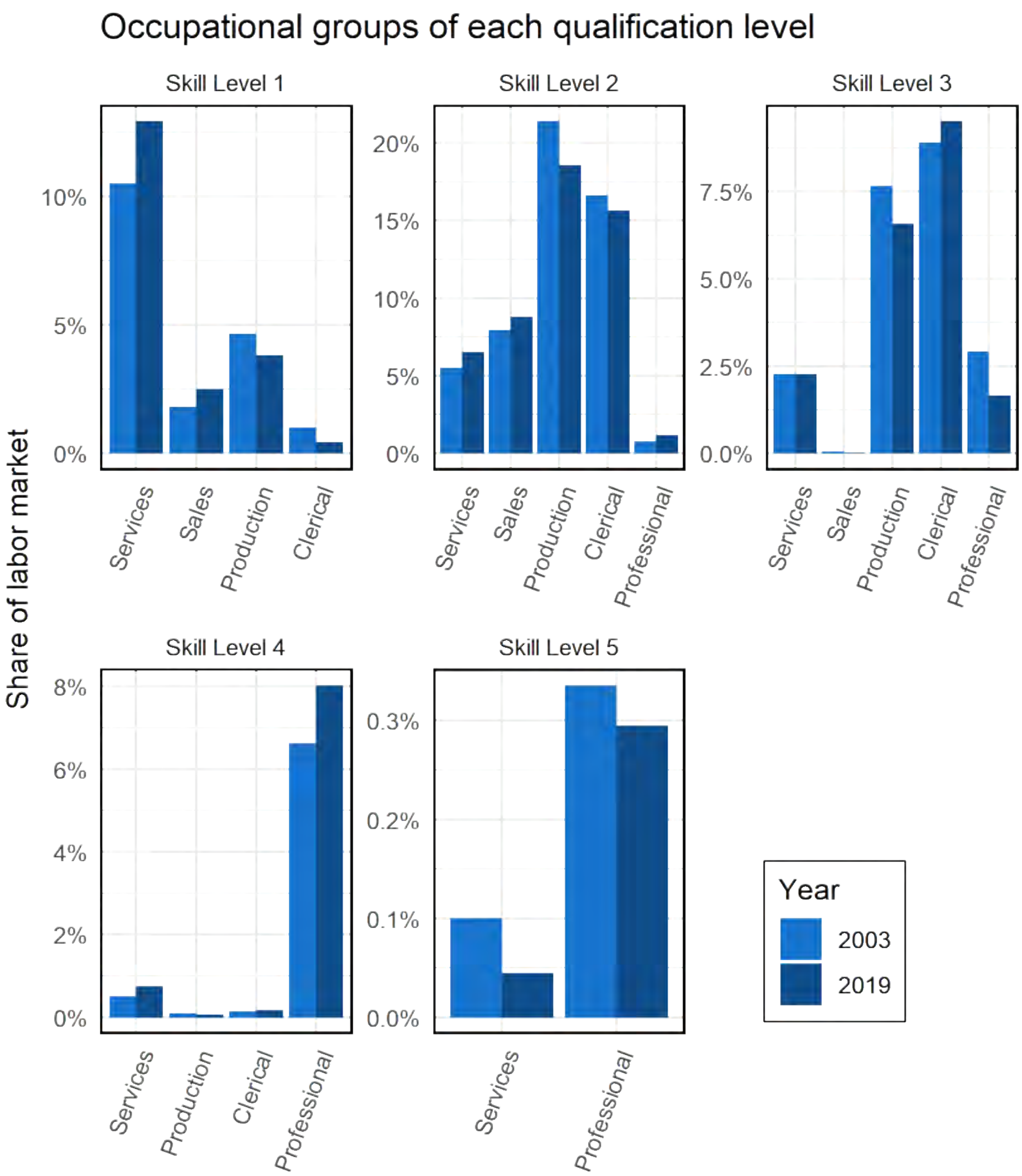


Figure 3 – Loess smoothing of the change in proportion of the labor market taken up by each occupation, weighted by the number of employees. Occupations were sorted by their mean wages in 2003, and had their number of workers divided by the cumulative sum of workers up to that wage level, then multiplied by 100.

explains why the skill level as a whole declined. At the third skill level, there was the added effect of the rise in clerical occupations, while professional jobs fell. The storyline flips for the top two qualification levels, where employment was subtly pushed upwards by rising professional occupations.



Occupational groups of each qualification level
Share of labor market
Skill Level 1
10%
5%
0%
Services
Sales
Production
Clerical
Skill Level 2
20%
15%
10%
5%
0%
Services
Sales
Production
Clerical
Professional
Skill Level 3
7.5%
5.0%
2.5%
0.0%
Services
Sales
Production
Clerical
Professional
Skill Level 4
8%
6%
4%
2%
0%
Services
Production
Clerical
Professional
Skill Level 5
0.3%
0.2%
0.1%
0.0%
Services
Professional
Year
2003
2019

Figure 4

1.3 Tasks

The path towards understanding this decline in middle-skilled occupations leads to the very core of the literature regarding tasks. To explain why the adoption of computers in the workplace favored highly educated workers, Autor, Levy, and Murnane (2003) introduced the key distinction between routine tasks, defined as those which can be encoded by a series of deterministic steps, and their non-routine counterparts. While high-skilled non-routine workers thrived following technology-induced productivity boosts, routine work began to be executed by computers.

It was then understood that technical change spurs a substitution of routine labor, most common for low-skilled occupations, as well as for middle-skilled clerical jobs, while high-skilled non-routine occupations, comprising managerial interpersonal tasks and analytic thinking, grow. On the other hand, low-skilled occupations in the service sector are often non-routine, involving customer service and complex manual tasks. As those also grow due to technical change, the negative effect on low-skill occupations is offset, more clearly outlining the reasoning behind job polarization.

In summary, the routinization hypothesis outlined by Acemoglu and Autor (2011) deems technical change to be not skill-biased, but rather routine-biased. As technical developments allow for the substitution of routine tasks by computers, middle-skill clerical occupations falter. The increased productivity granted by technology complements high-skilled workers, who experience wage gains, while not as widely replaceable low-skill service occupations also benefit from higher demand.

The next chapter will concern applying these task-based notions to the Brazilian labor market.

2 The skill profile of Brazilian occupations

2.1 Data Sources

Data characterizing each occupation is taken from the Brazilian Qualifications Framework, *Quadro Brasileiro de Qualificações* (QBQ). The QBQ maps each occupation, encoded by its 6-digit code in the *Classificação Brasileira de Ocupações* (CBO), the Brazilian Occupation Classification, to a qualification score. Furthermore, this qualification score is built upon a multitude of parameters, detailing the types of skill, knowledge, and attitudes required in a given occupation.

As the QBQ is still in construction, only 1800 out of the 2734 unique CBO occupation codes found in the data for this period have been described. Those, however, make up over 80% of the workforce. Figure 5 shows the distribution of workers across the CBO great groups – given by the code's first digit – and specifies how many of them are in occupations included in the QBQ. The coverage is notably not homogeneous, which stems from a conscious decision by the QBQ to track low-skill occupations first. All scores are given in a discrete scale, ranging from 1 to 5, and which will be extended to a maximum of 8 as more qualified occupations are tracked.

The QBQ microdata is split into three databases, each listing, respectively, the skills, types of knowledge, and attitudes present in an occupation. Each characteristic is also paired along with frequency, importance, and depth scores, in the same discrete scale. There are 160 different individual skills, over 1000 types of knowledge, and 13 types of attitude.

All of those variables are structured into groups of different aggregation levels. For example, skills are sorted into the 3 domains of Cognitive, Practical, and Physical, psychomotor and sensory. Those are further broken down: Cognitive skills are made up of categories such as communication and working with numbers. Each category then contains a series of individual skills.

Job market data comes from the *Relação Anual de Informações Sociais* (RAIS), a longitudinal matched employee-employer dataset from the Ministry of the Economy tracking virtually all formal job links in Brazil. Each link accompanies firm-level data, such as a firm identifier and sector code, worker-level data, such as an individual identifier across the years, education level and age, as well as data particular to each link, most notably the wage and the occupation code, which is also encoded by the CBO. Because of this, every employer-employee

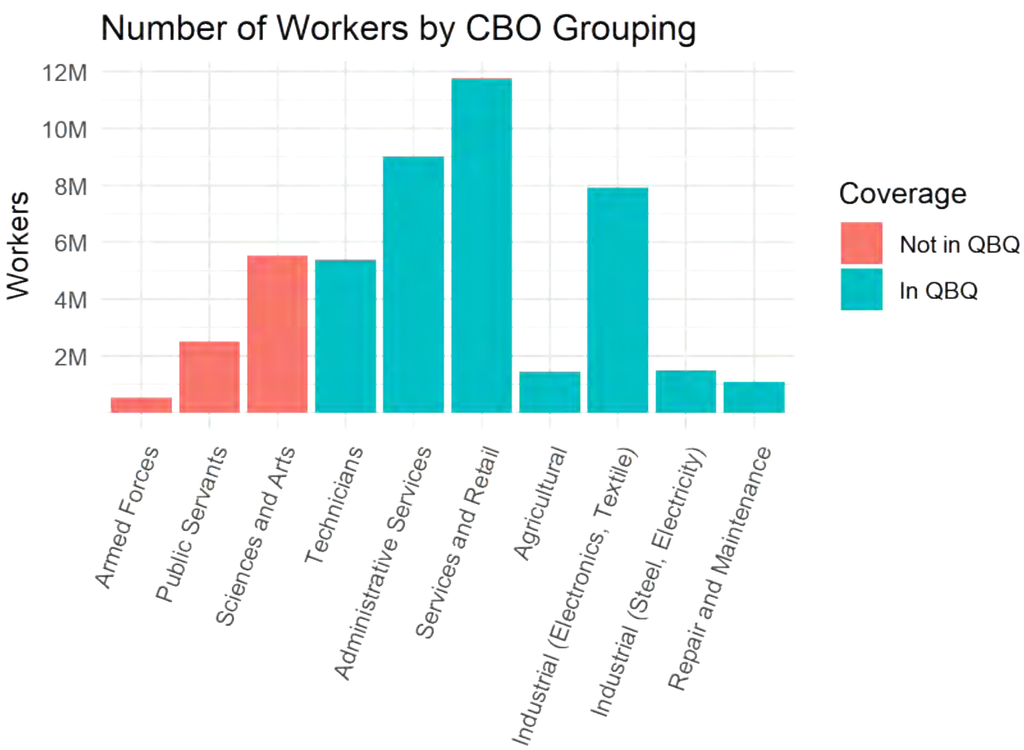


Figure 5

link can be matched with the vast QBQ occupational data.

2.2 Measuring routine tasks

The QBQ lists one particular skill named "applying simple and routine instructions", whose importance score, between 1 and 5, serves as a measure of routineness. When this skill is absent, I consider the score to be zero.

The top panel of Figure 6 shows how the average routine score, ranging from 0 to 5, is highest for low-skill occupations and goes down nearly monotonically as skill increases. The panel on the right plots the proportion of workers whose occupation's routine score is 4 of 5, described as a "large" or "very large" importance by the QBQ. Over 80% of workers in the lowest qualification level work in routine-intensive occupations by this metric.

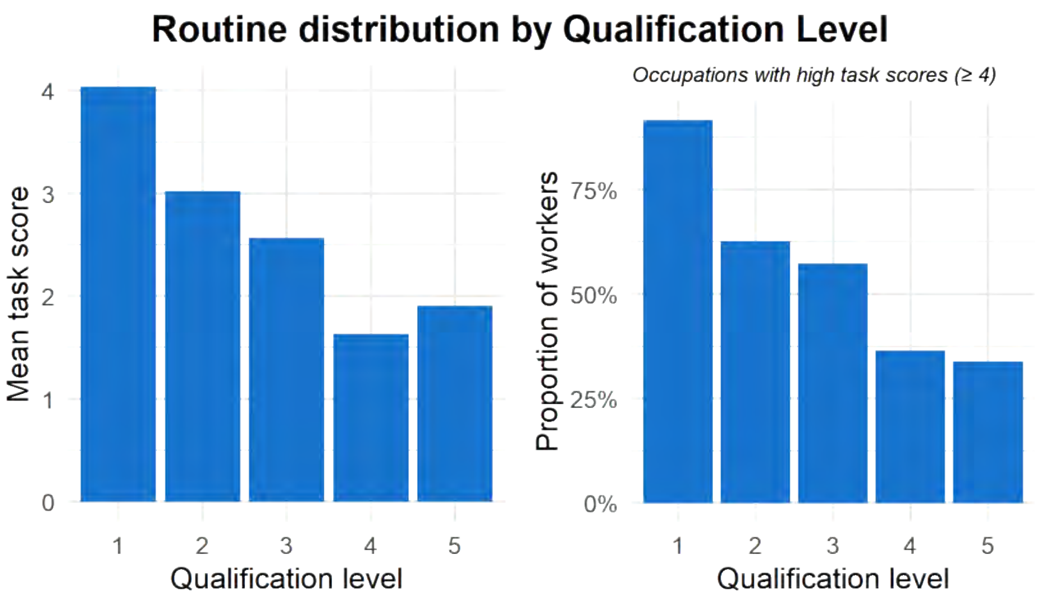


Figure 6

Substituting education groups for the qualification levels, Figure 7 shows an analogous format, where higher education corresponds to lower routineness. It is however noteworthy that, among the occupations so far covered by the QBQ, over 70% of workers with no High School degree work in highly routine occupation, which falls to 60% for High School educated workers, and still remains above 50% for workers with a College degree.

Comparing between occupational categories, Figure 8 shows that the often low-skilled service occupations are highly routine, while the also low-skilled sales occupations are the least routine of all groups. Middle-skilled production and clerical occupations are also mostly routine, and high-skilled professional occupations are mostly non-routine.

With these events in mind, I now look at the changes in this routine measure, and use them to interpret the overarching movements seen in section 1.2. In

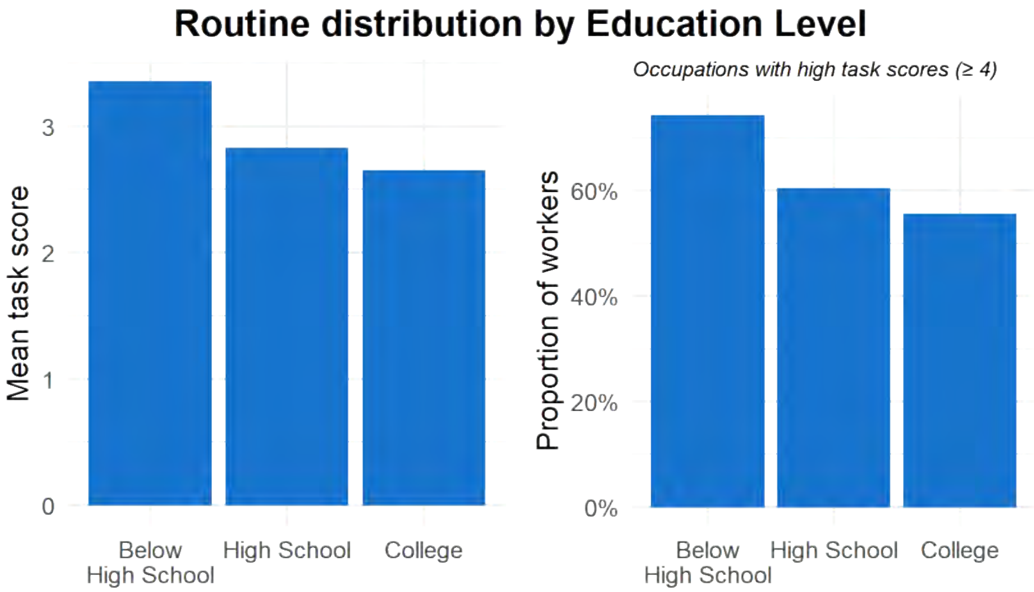


Figure 7

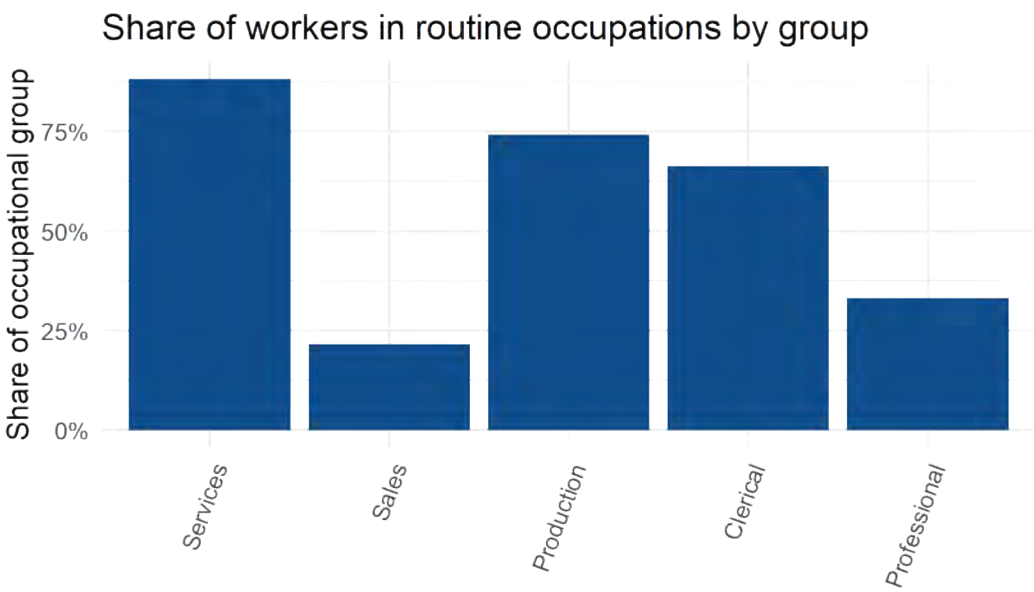


Figure 8

Figure 9 are the changes in the average routine level of workers, both by qualification levels and by education levels. In a puzzling departure from theoretical predictions, the left-most panel shows that, within low- and middle-skill occupations, work has mostly shifted in favor of routine occupations. A steep decline in routineness is only present at the highest qualification levels. This coincides with what is seen when analysing changes within education groups, where only college-educated workers have strayed from routine occupations.

In the light of Figure 8, it becomes apparent that although the lower skill levels underwent a large shift of declining production (and sometimes clerical) jobs, whose place was occupied by service jobs, these growing service jobs are actually

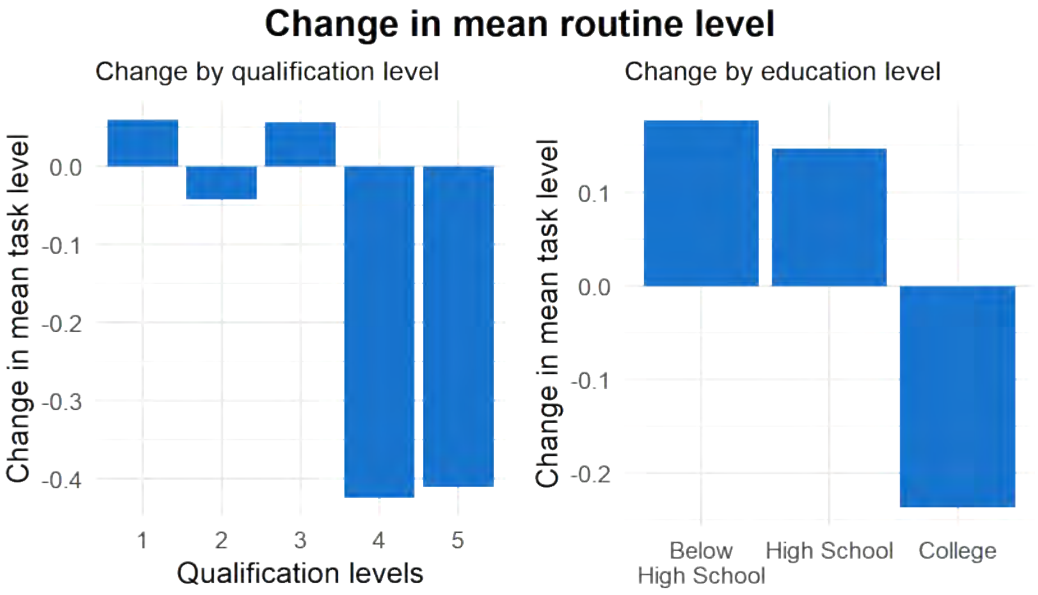


Figure 9

also highly routine. Therefore, there was no large net change in routineness at those levels, despite a large-scale change in the type of job being performed. Meanwhile, the movement of more workers towards high-skilled professional occupations led to a sharp fall of routineness at those levels, and for college-educated workers.

As a whole, this indicates that, although routineness has faced a very slight decline overall – from an average of 2.08 in 2003 to 2.04 in 2019 – its shifts do not explain the larger scale occupational changes that have taken place in Brazil. In particular, there has been no clear decline in middle-skill routine occupations that would have contributed to a potential pattern of job polarization.

Besides the routine index, I constructed indices for autonomy, abstract reasoning, interaction, and technology. As they were build from groups of hand-picked variables, the way to calculate them is not as clear cut. Most importantly, they add little new information when compared to the more robust PCA indices built in the next section, and were therefore relegated to Appendix A.

2.3 Principal Component Analysis

Thus far, I have taken the more conventional approach to task measures, which consists of selecting a a key variable – or a group of variables – to represent a task, and using them to compute simple means and form indices. In this chapter, I instead employ statistical methods designed for dealing with high dimension – fitting for the hundreds of QBQ variables. This has a couple of advantages: first, they allow us to let the data speak for itself. With minimal human manipulation and no manual selection of variables, these methods will bring to light the underlying patterns and what the actual main correlations between tasks are. Secondly, the volume of data is simply way too large in the context of the QBQ, so a manual approach is sure to ignore relevant variables and not take full advantage of the information contained in the data. On the downside, these methods are purely statistical, as opposed to the more manual method, which is backed up by economic intuition. Therefore, the way these algorithms summarise the data may not always be intuitive, meaning that effort is required to interpret the results and reach tangible conclusions.

The core idea of Principal Component Analysis (PCA) is that, when many variables interact, their movement can often be summarised by a few hidden factors. In the case of the QBQ, skills such as “Problem solving”, “Applying high-complexity scientific principles”, and “Data analysis” are likely all reflecting of some underlying analytic ability. Because of this, PCA is able to explain most of the variance of the dataset with only a few summarising variables, the so-called principal components.

These components are simply some linear combination of the variables, whose coefficients – called loadings – solve the optimization problem of explaining as much variance as possible. Once the factor loadings have been estimated, I compute the value that the principal components attain for each occupation. This is done by plugging in the realized values of the variables into the linear combination whose coefficients are the factor loadings. The resulting values are named factor scores. The factor scores obtained in this section will be the main variables used to estimate returns to skills and to skill switches in the next chapter.

2.3.1 Previous uses in the literature

Principal Component Analysis, as other statistical methods, sets up a delicate trade-off between interpretability and accuracy of the estimates. As a result, how exactly to apply this method is not set in stone.

In the task literature, Poletaev and Robinson (2008) introduced a novelty from other areas by employing PCA as an all-purpose tool to measure each occupation’s task profile. With no arbitrary manipulations of variables, they extracted a few principal components from their whole set of skill variables. Each component was taken as measure of some aspect of skill, and the values of the loadings allowed them to attibute an *a posteriori* interpretation to each component.

Applying this method to the QBQ data allowed for extremely limited interpretability. The first principal component of all skills dominates over half the variance, being largely determined by physical skills. For the subsequent components, the distribution of loadings does not match up with any QBQ skill groups. Using extensions to PCA such as Sparse PCA and using a polychoric covariance matrix (since the data are discrete) made small differences in the results.

I now turn to a method that forces some interpretability out of PCA estimations. It consists on first sorting variables into interpretable groups of skills, and then executing PCA separately on each. This has been used by Autor, Levy, and Murnane (2003), and later by Goos, Manning, and Salomons (2011), although solely for the purpose of robustness checks. It was a tool to verify whether their results would be largely affected by constructing of indices through PCA rather than by their manual assignment of variables.

I employ this method on the QBQ data by taking advantage of their preset skill categories, by which each skill is considered Cognitive, Physical (along with Psychomotor and Sensory), or Practical. I extract the first Principal Component of the variables in each group and take it to be the index for how Cognitive, Physical, or Practical a given occupation is. To reduce noise, I keep only the skills present in at least 20 occupations.

At the cost of some interpretability – it is not quite as straightforward as taking means of small groups of variables – the principal components are always continuous variables built from many skill scores, which should make them more robust, and less prone to distortions due to arbitrary variable choices.

Descriptive statistics resulting from this application of PCA are the topic of subsection 2.3.2.

2.3.2 Descriptive statistics for the skill indices

Before leveraging these Principal Components as explanatory variables of wage regressions, it is useful to get a sense for what information their measurements bring to the table. In general, it is expected that cognitive skills be higher for higher-skilled occupations, whereas physical skills are expected to be lower. Practical skills, however, are more of a mixed bag since they encompass both higher-skilled interpersonal and organizational skills, as well as more routine operational skills.

To show how these indices are actually composed, Figure 10 plots the PCA loadings of each skill making up the cognitive, physical, and the practical indices, with labels below for the subgroups the skills belong to. Each tile stands for an individual skill, with a blue color meaning that its PCA loading is positive – contributing positively to the overall index. They can be thought of as a rescaled weighted average, where few variables take on negative loadings – which would mean that they behave against the remainder of the pack. This indicates that the skill groups are cohesive, and thus capable of forming informative indices. It is of note that the cognitive index relies quite heavily on "communication" and "working with numbers" skills, the physical index is mostly made up of "physical" variables, which are closest to constituting manual labor, and the practical index is mostly "interpersonal" and "other practical".

Figure 11 plots the mean value of the PCA scores for each education level. The scores were standardized to a mean of 0 and a standard deviation of 1. As expected, cognitive skills increase with schooling, while physical skills decrease. Practical skills are also higher for higher levels of education.

To verify whether the indices line up with the intuition for different types of jobs, Figure 12 shows the mean PCA scores of each ISCO occupational group, along with the previous Routine index, also standardized. Services and sales occupations are lowest in cognitive skills, which are slightly higher in production occupations, and increase drastically for the higher-skilled clerical and professional occupations. Meanwhile, physical skills are highest for both services and production occupations. Practical skills are highest for clerical and professional occupations, but remain mostly constant across all groups.

Figure 13 shows how the PCA scores are distributed by income levels. While cognitive skills are increasing in wages, and physical skills are decreasing, practical skills are overall increasing, but peak around middle-skilled occupations.

To see how these scores have changed, Figure 14 plots the mean PCA scores by year. Those with no High School degree have moved towards slightly more physical, less cognitive, and less practical occupations. High school graduates faced the highest level of change, also with an increase in physical skills along with a decrease in cognitive and practical skills. Note that the overall scores were standardized to mean zero in 2019, which is why workers with High School education, that make up most of the labor force, converge to values close to zero. College graduates show the same overall pattern, but at a smaller scale.

Although all tasks follow some previously discussed patterns, Table 1 shows that each skill index brings some new independent dimension to the table. Since the routine skill has very little weight in the Practical index in favor of interpersonal skills, the Routine and Practical indices are barely correlated. As expected, Cognitive skills are negatively correlated to Routine and Physical ones, as they're on opposite ends of the skill spectrum.

Although it allows for more robustness, PCA still faces the limitation that the QBQ data is static, giving fixed skill levels to each occupation. In turn, it cannot capture within-occupation task changes. Only the movement of workers

Table 1 – Correlation matrix between skill indices

	Routine	Cognitive	Practical	Physical
Routine	1.00			
Cognitive	-0.40	1.00		
Practical	-0.09	0.52	1.00	
Physical	0.23	-0.38	-0.39	1

Variables were standardized and covariances were calculated weighting by the number of employees in each occupation in 2019.

between occupations causes the mean value of each score to change over time.

The next chapter will apply these PCA scores to an analysis of wage returns, obtaining variation in tasks from the displacement of workers. To capture some level of within-occupation variance, I will also use each worker's cumulative experience in a skill as an explanatory variables, which can vary wildly even in a given occupation.

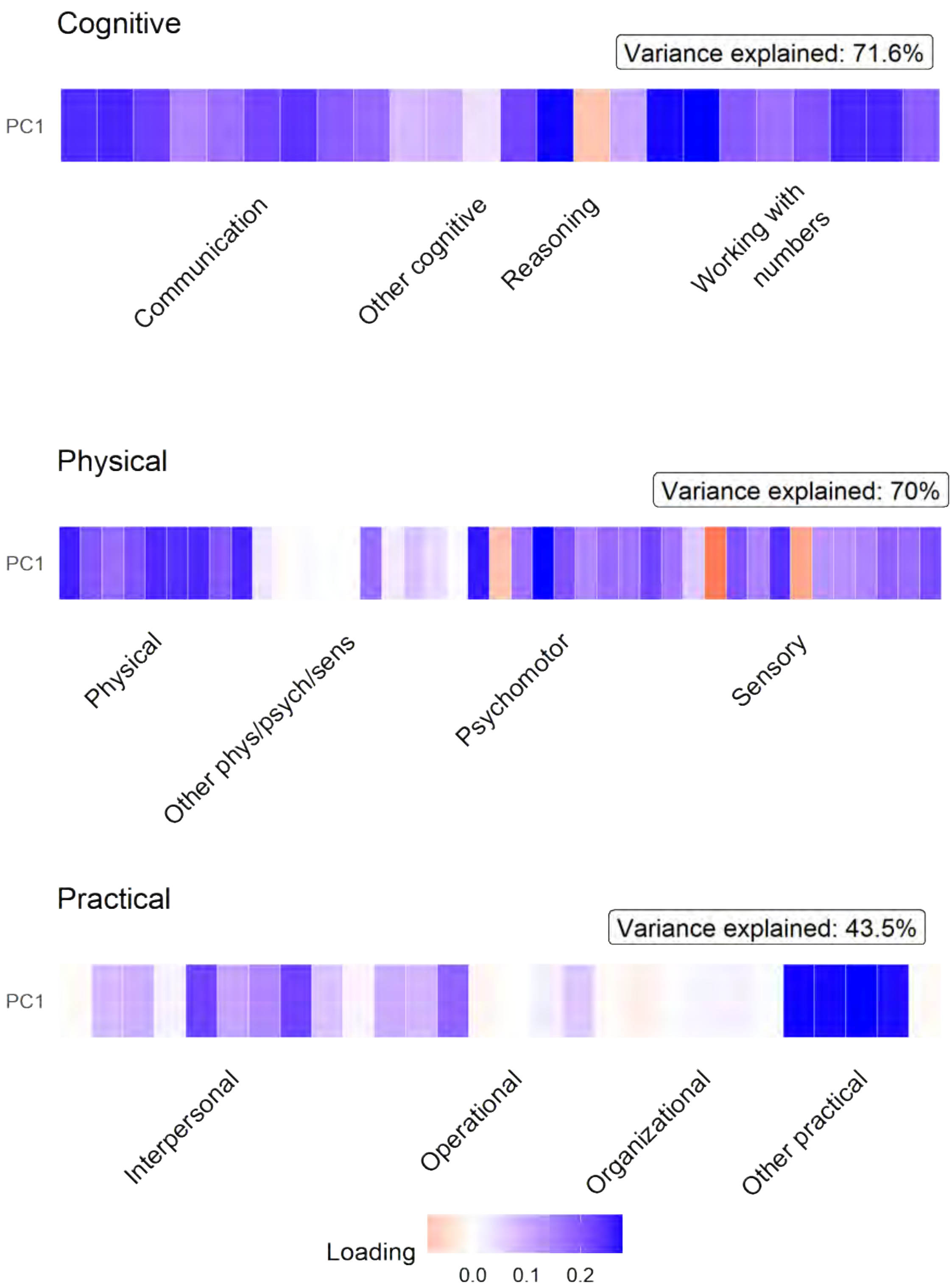

Cognitive
Variance explained: 71.6%
PC1
Communication
Other cognitive
Reasoning
Working with numbers
Physical
Variance explained: 70%
PC1
Physical
Other phys/psych/sens
Psychomotor
Sensory
Practical
Variance explained: 43.5%
PC1
Interpersonal
Operational
Organizational
Other practical
Loading
0.0
0.1
0.2

Figure 10

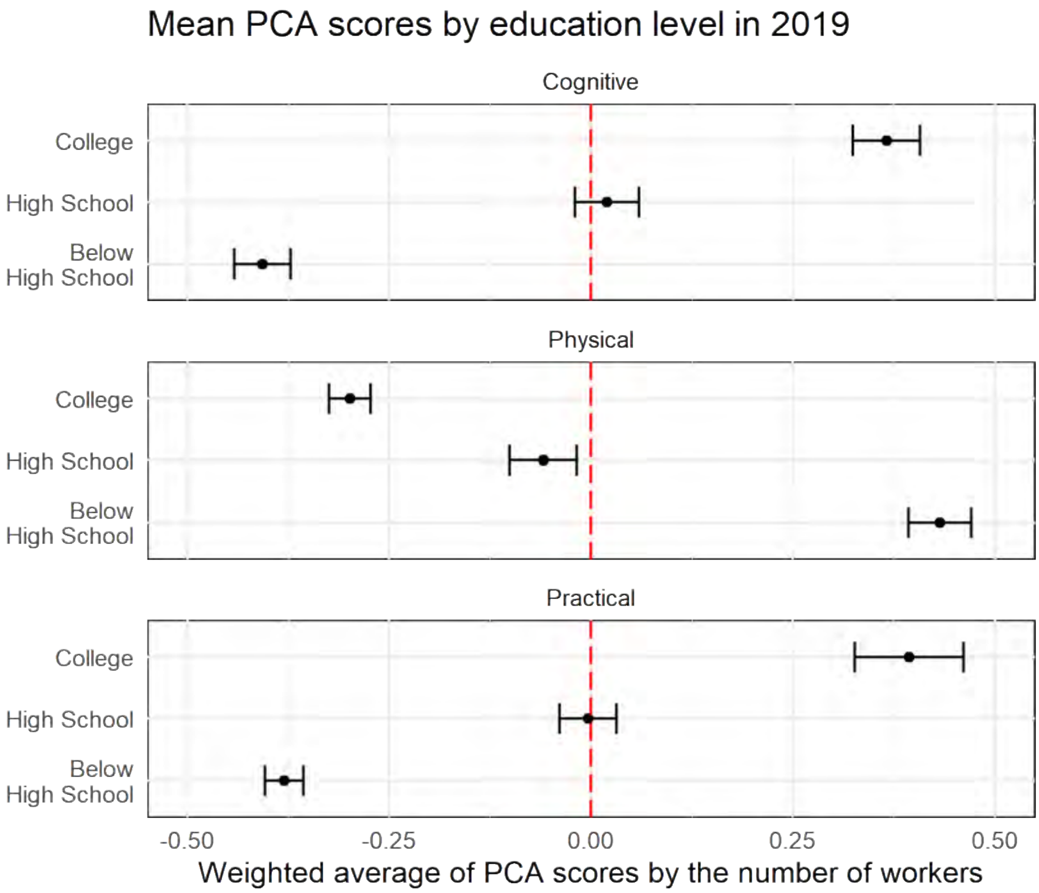


Figure 11

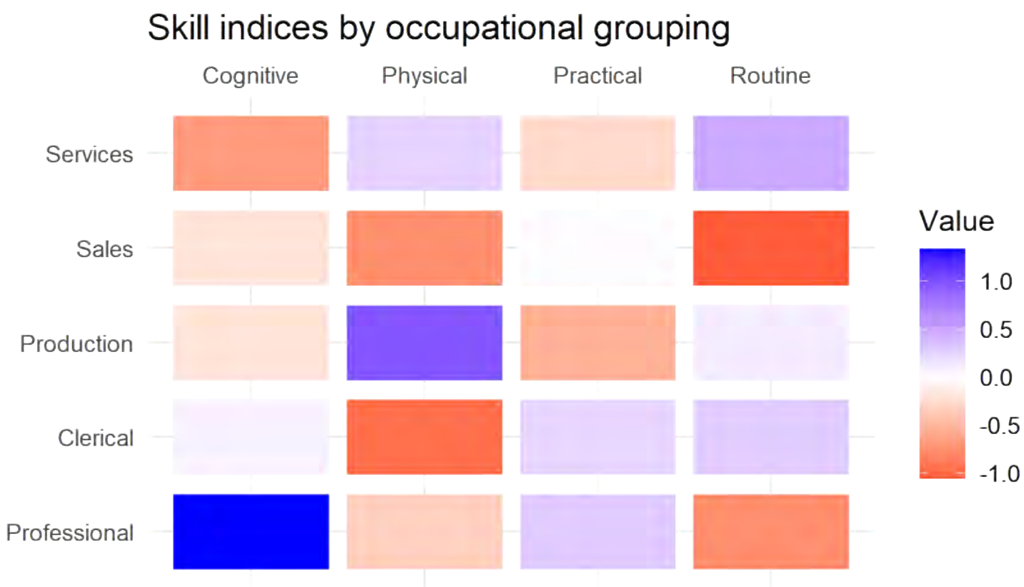


Figure 12

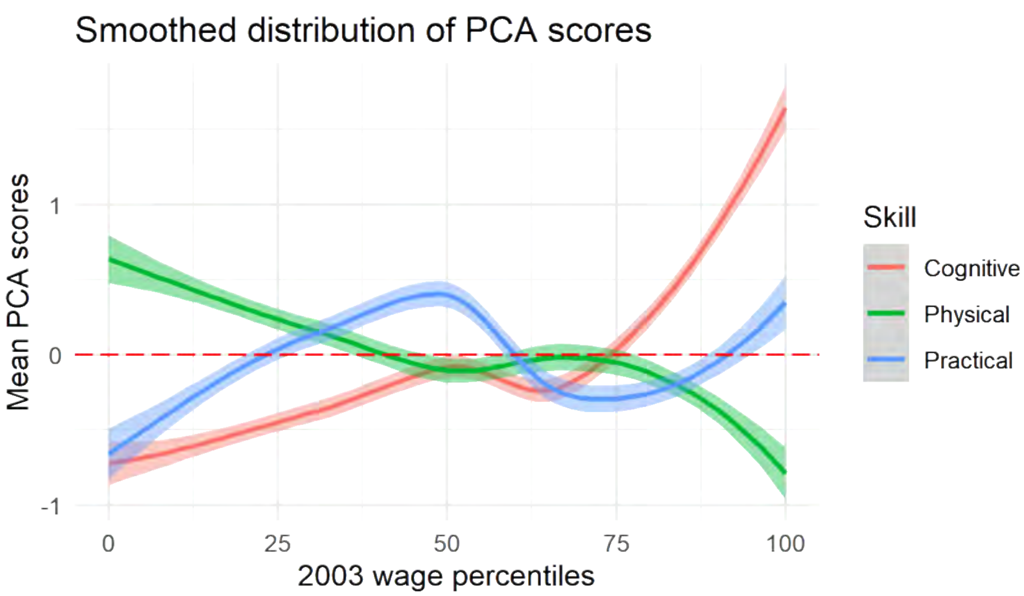


Figure 13

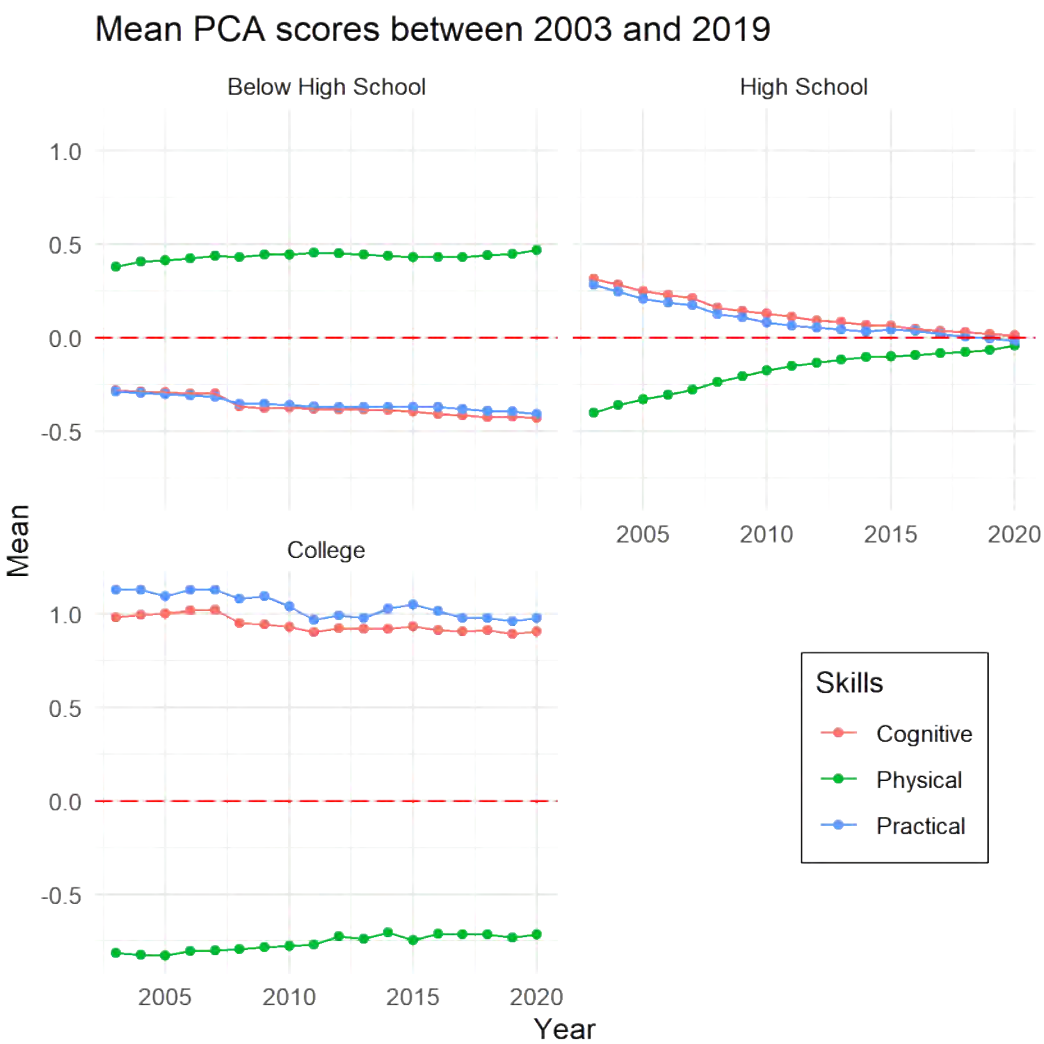


Figure 14

3 Empirical strategy

The previous chapters have looked at the task composition of the Brazilian job market, broken down by education and income levels. In this chapter, I estimate (i) how much each task is rewarded in the labor market, and (ii) how much transitioning into occupations with different skill sets can impact earnings.

3.1 Returns to skills

To measure the wage returns to each skill, I estimate a simple wage regression, with skill indices as the main covariates. In this chapter, I use both the routine index from section 2.2, which is standardized to line up with the other skill variables, and the three PCA skill measures from section 2.3. Equation (1) models the wage returns to each Skill^j, where $j = 1, \ldots, 4$ indexes Routine, Cognitive, Physical and Practical scores respectively.

$$\log(w_{it}) = \beta_0 + \sum_{j=1}^{4} \beta_j \, \text{Skill}_i^j + \theta_t + \gamma X_{it} + u_{it}. \tag{1}$$

For each individual i at year t, w_{it} is the average monthly wage, θ_t represents a series of year dummy variables, and X_{it} is a vector of controls varying by specification. I estimate three different specifications for (1): specification (i) only has dummies for each year and state; (ii) adds individual-level controls for education, age, tenure in the current job, gender, race, and disability; and (iii) adds dummies for each industry. Since skills are observed at the occupation level, I am unable to control for occupation in this model.

Not controlling for individual workers across time is a source of bias, as there will always be confounding factors that drive some workers into particular occupations instead of others: employees with higher cognitive skill are both more likely to work in cognitive-intensive occupations and to have other desirable characteristics in terms of aptitude and learning on the job. Is is important to draw a distinction between the tasks observed in occupations and the intrinsic skill endowments of each individual, which are not observable. Autor (2013) explains that the tasks performed in practice are not only a function of each worker's skills and preferences, but also of the job market wage incentives themselves. As a result, observed tasks are always endogenous variables, with workers self-selecting into performing tasks that are valued more highly, even if the higher pay is not actually caused by their skills.

Equation (2) mitigates the former source of bias, by including fixed effects for each individual to control for their unobservable characteristics which are constant over time. Fixed effects for each year are given by θ_t, and fixed effects for each worker are given by μ_i. As most individual-level controls are fixed over time, I include only state and industry fixed effects the vector of controls X_{it}.

$$\log(w_{it}) = \beta_0 + \sum_{j=1}^{4} \beta_j \,\mathrm{Skill}_i^j + \theta_t + \mu_i + \gamma X_{it} + u_{it} \tag{2}$$

To investigate whether these wage returns are heterogeneous across education levels, I estimate (1) separately for three subsamples of workers with different education levels. I measure separate returns to each skill for workers below high school, with a high school degree, and for college graduates. In this case, X_{it} contains all controls available, mimicking specification (iii).

3.2 Skill experience

The largest remaining challenge to estimating the effect of tasks on wages is that tasks are treated as constant for all individuals in an occupation, across all years. This clashes with the finding by Atalay et al. (2020) that tasks are highly heterogeneous across workers in the same occupation. If there are workers within the same occupations who perform more cognitive tasks, and who also earn more as a result, the magnitude of wage returns by task from (1) and (2) will be underestimated, as all this within-occupation variance is unobserved.

In an attempt to extract this within-occupation variance, I follow Gonzaga and Guanziroli (2019) to construct a time-varying measure of accumulated task experience for each individual, which serves as an approximation of each worker's built-up specific human capital in each skill. These measures are the sum of all skill scores throughout a worker's career up to that point, weighted by the time spent at each job, namely

$$\mathrm{Skill}_{it}^{\mathrm{exp}} = \frac{1}{12} \sum_{\tau \leq t} \mathrm{Skill}_j \times \text{months in occupation } j \text{ in year } \tau,$$

where worker i earned most of their income in year τ from occupation j.

I then estimate (1) again, but replacing the previous current task variables with their task experience counterparts. As this approach allows for within-occupation variance in skills, I also control for 6-digit occupations in the last specification. These estimates measure how much within-occupation discrepancies in skills can account for differences in pay.

3.3 Occupational distance and wages

I now measure the effect of moving to an occupation requiring different skills on wages, regardless of which skills they are. According to Poletaev and Robinson (2008), shifting towards occupations further away in the skill spectrum leads to losses of specific human capital, not rewarded in the new occupation. While a worker may earn a wage premium for beginning to perform certain tasks, they may also suffer losses from the lack of experience in that task.

I calculate the distance between two occupations following Poletaev and Robinson (2008), where it is simply the euclidean distance between the skill vectors for each occupation, which in my case I take to be

$$(\mathrm{Skill}_i^j)_{j=1}^4 = (\mathrm{Cognitive}_i, \mathrm{Physical}_i, \mathrm{Practical}_i, \mathrm{Routine}_i).$$

The distance between occupations a and b is then computed as

$$D_{a,b} = \sqrt{\sum_{j=1}^{4} \left(\mathrm{Skill}_a^j - \mathrm{Skill}_b^j\right)^2}.$$

I use this distance measure to estimate the model

$$\log(w_{it}) = \beta_0 + \beta\, D_{it} + \theta_t + \gamma X_{it} + u_{it}, \tag{3}$$

where D_{it} is the distance between a worker's current and previous occupations, and the vector X_{it} of controls follows specifications (i) and (iii) as before, with a third also including the four skill indices as explanatory variables.

I once again estimate (3) separately for workers of different education levels. If occupation distance affects wage changes through the loss of specific human capital, its effects should be different for workers of varying levels of schooling, a measure of general human capital.

4 Results

4.1 Returns to skills

The main results are presented in Table 2 from equations (1) and (2), with the log average monthly wage as the dependent variable and the four standardized skill indices as the main regressors. The first three columns contain coefficients from specifications with increasing controls, while the last one is estimated with fixed effects at the individual level. Across all specifications, the routine index is associated with lower wages, with workers performing one additional standard deviation of routine work earning between 2-3% less. Cognitive skills, on the other hand, are linked to substantially higher wages, over 5% higher per standard deviation in the most conservative estimates. Surprisingly, physical skills have positive coefficients, while practical skills have negative ones. As will be shown in Table 3, however, these effects are not consistent across education levels.

Table 2 – Wage returns to skills

	Dependent variable: Log wage			
	(1)	(2)	(3)	(4)
Routine	-0.025***	-0.032***	-0.035***	-0.021***
	(0.0001)	(0.0001)	(0.0001)	(0.0003)
Cognitive	0.217***	0.143***	0.119***	0.056***
	(0.0002)	(0.0001)	(0.0001)	(0.0004)
Physical	0.037***	0.037***	0.027***	0.023***
	(0.0001)	(0.0001)	(0.0001)	(0.0004)
Practical	-0.029***	-0.074***	-0.040***	-0.021***
	(0.0002)	(0.0002)	(0.0002)	(0.0005)
N	25,524,013	22,196,751	22,196,706	25,523,963
R-squared	0.20	0.44	0.53	0.42
Year and state FEs	Yes	Yes	Yes	Yes
Schooling, tenure, age, race, gender, disability	No	Yes	Yes	No
Industry FEs	No	No	Yes	Yes
Individual FEs	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Performing separate estimations of wage returns for each education level, Table 3 highlights the heterogeneity of the aforementioned returns to skills. These follow the specification in the third column of Table 2, with the most controls, but each column shows the estimates obtained when limiting the sample to only one education level. The first row shows that routine work is penalized at all education levels, but increasingly so for higher education levels. Cognitive skills are rewarded more intensely for higher educated workers. Physical skills, while linked to positive wage premia for those without a high school degree of with high school education, have considerable wage penalties for college-educated workers. In an analogous pattern, practical skills have a positive coefficient for the subsample of workers below high school, which turns negative for the higher educated.

Table 3 – Wage returns to skills by education level

	(1) Below High School	(2) High School	(3) College
Routine	-0.032***	-0.039***	-0.045***
	(0.0002)	(0.0001)	(0.0006)
Cognitive	0.096***	0.115***	0.141***
	(0.0002)	(0.0002)	(0.0006)
Physical	0.067***	0.013***	-0.099***
	(0.0002)	(0.0002)	(0.0009)
Practical	0.031***	-0.042***	-0.089***
	(0.0003)	(0.0002)	(0.0005)
N	9,680,284	11,032,288	1,484,134
R-Squared	0.46	0.50	0.47

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2 Skill experience

Table 4 displays the results of section 3.2, where the previous skill indices are replaced by the cumulative experience held by workers in each skill. The first column estimates the same specification of the third column of Table 2, with many controls but no individual- or occupation-level fixed effects. These results are extremely close to the ones found for the skill variables themselves, indicating that both current skill level and skill experience have similar and confounding returns. Even the second column, with individual fixed effects, does not stray far from the previous results. The third column takes advantage of the fact that skill experience, unlike current skill levels, varies between workers in the same occupation. I find that, among workers in the same 6-digit occupation, those with more previous experience in cognitive-intensive occupations earn more, while those with more physical experience earn less. The coefficients for routine and practical tasks are also significant but of small magnitude.

Table 4 – Wage returns to skill experience

	Dependent variable: Log wage		
	(1)	(2)	(3)
Routine Exp.	-0.028***	-0.009***	-0.003***
	(0.0001)	(0.0002)	(0.0002)
Cognitive Exp.	0.097***	0.023***	0.026***
	(0.0001)	(0.0002)	(0.0002)
Physical Exp.	0.017***	-0.004***	-0.017***
	(0.0001)	(0.0002)	(0.0002)
Practical Exp.	-0.033***	-0.005***	-0.006***
	(0.0001)	(0.0002)	(0.0003)
N	22,196,706	25,524,009	22,196,706
R-squared	0.53	0.38	0.59
Year, state, and industry FEs	Yes	Yes	Yes
Schooling, tenure, age, race, gender, disability	Yes	No	Yes
Individual FEs	No	Yes	No
6-digit occupation FEs	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 Occupational distance and wages

Table 5 presents the results of (3), showing how, upon a job switch, a worker's wage changes according to the distance between the previous and current occupation. The negative coefficients in the first row align with the idea from Poletaev and Robinson (2008) that bigger leaps between occupations are associated to smaller wage gains, with other factors controlled for. The third column shows that these results are consistent even when controlling for the skills of the current occupation themselves, and also verifies that the returns to each skill maintain the signs found previously for the case of job switches, and controlling for occupational distance.

In Table 6, the previous analysis is broken down by education level. The first row, which shows the coefficient for the variable of interest in this case, indicates that the penalty for making higher-distance occupational switches is steeper for college-educated workers. As those are the workers with the most build-up of human capital, this result supports the hypothesis by Poletaev and Robinson

Table 5 – Wage returns to job switches

	Dependent variable: Log wage		
	(1)	(2)	(3)
Occ. Distance	-0.049***	-0.022***	-0.031***
	(0.0002)	(0.0002)	(0.0002)
Routine			-0.044***
			(0.0002)
Cognitive			0.133***
			(0.0003)
Physical			0.011***
			(0.0002)
Practical			-0.071***
			(0.0003)
N	13,647,680	10,192,313	10,192,313
R-squared	0.05	0.37	0.41
Year and state FEs	Yes	Yes	Yes
Schooling, tenure, age, race, gender, disability	No	Yes	Yes

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(2008) that such wage losses are due to a loss of specific human capital. The results to skills are still similar to the ones seen in Table 3.

Table 6 – Wage returns to job switches by education level

	(1) Below High School	(2) High School	(3) College
Occ. Distance	-0.019***	-0.030***	-0.043***
	(0.0003)	(0.0002)	(0.0004)
Routine	-0.035***	-0.051***	-0.031***
	(0.0003)	(0.0002)	(0.0006)
Cognitive	0.116***	0.129***	0.140***
	(0.0004)	(0.0004)	(0.0008)
Physical	0.042***	0.015***	-0.080***
	(0.0003)	(0.0003)	(0.0011)
Practical	0.052***	-0.060***	-0.112***
	(0.0007)	(0.0004)	(0.0006)
N	4,021,810	4,651,581	1,518,922
R-Squared	0.37	0.38	0.39

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Conclusion

I find evidence that the skill composition of occupations is relevant to explain wage discrepancies, even when controlling for observables such as education level and sector. Workers in occupations with a more cognitive skill profile are consistently rewarded more than their routine counterparts. I also observe that these effects are heterogeneous across education levels, with workers with less schooling having a smaller magnitude of wage gains and losses for cognitive and routine skills respectively. They are also able to gain higher wages from higher physical and practical skill levels, which have the reverse effect for the higher-educated.

An analysis of the accumulated skills along each worker's career reveals that, even for workers in the same occupation, cognitive skills are strongly associated to wage gains, while physical skills are associated to wage losses. When workers switch occupations, I find skill profiles to still have significant returns, but also see negative wage changes due to higher occupational distances. This matches Poletaev and Robinson (2008), where workers incur in specific human capital losses when switching to occupations further away in the skill spectrum. This impact of occupational distance is more pronounced for workers with a college degree.

Alongside these main empirical findings, I outline a method to obtain aggregate skill measures from the QBQ. This database is very recent – not yet even fully completed – and poses particular challenges due to the wide array of variables and level of detail it presents. This analysis could be extended to different dimensions of occupation characteristics, taking advantage not only of the skills database, but also of the knowledge and attitudes ones. One could take a more fine-tuned approach to measure some narrower type of skill, as I explored briefly in Appendix A. The knowledge database is the vastest part of the QBQ, with over 1500 individual entries in over 200 distinct fields. Since it is so detailed, it makes it possible to track particular types of computer technologies used in each occupation, for example. In summary, the QBQ has plenty of untapped potential to increase our understanding of how different facets of occupations interact and affect workers.

Bibliography

[1] Daron Acemoglu and David Autor. "Skills, tasks and technologies: Implications for employment and earnings". In: *Handbook of labor economics*. Vol. 4. Elsevier, 2011, pp. 1043–1171.

[2] Enghin Atalay et al. "The evolution of work in the United States". In: *American Economic Journal: Applied Economics* 12.2 (2020), pp. 1–34.

[3] David H. Autor. "The "task approach" to labor markets: an overview". In: *Journal for Labour Market Research* 46.3 (Sept. 2013), pp. 185–199. ISSN: 1867-8343. DOI: 10.1007/s12651-013-0128-z. URL: https://doi.org/10.1007/s12651-013-0128-z.

[4] David H. Autor and David Dorn. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market". In: *American Economic Review* 103.5 (Aug. 2013), pp. 1553–97. DOI: 10.1257/aer.103.5.1553. URL: https://www.aeaweb.org/articles?id=10.1257/aer.103.5.1553.

[5] David H. Autor and Michael J. Handel. "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages". In: *Journal of Labor Economics* 31.S1 (2013), S59–S96. DOI: 10.1086/669332. eprint: https://doi.org/10.1086/669332. URL: https://doi.org/10.1086/669332.

[6] David H. Autor, Frank Levy, and Richard J. Murnane. "The Skill Content of Recent Technological Change: An Empirical Exploration". In: *The Quarterly Journal of Economics* 118.4 (2003), pp. 1279–1333. ISSN: 00335533, 15314650. URL: http://www.jstor.org/stable/25053940 (visited on 11/19/2022).

[7] Carlos Henrique Corseuil, Miguel N. Foguel, and Gustavo Gonzaga. "Apprenticeship as a stepping stone to better jobs: Evidence from Brazilian matched employer-employee data". In: *Labour Economics* 57 (2019), pp. 177–194. ISSN: 0927-5371. DOI: https://doi.org/10.1016/j.labeco.2019.02.002. URL: https://www.sciencedirect.com/science/article/pii/S0927537119300089.

[8] David J Deming. *The Growing Importance of Decision-Making on the Job*. Tech. rep. National Bureau of Economic Research, 2021.

[9] Gustavo Gonzaga and Tomás Guanziroli. "Returns to experience across tasks: evidence from Brazil". In: *Applied Economics Letters* 26.20 (2019), pp. 1718–1723. DOI: 10.1080/13504851.2019.1593927. eprint: https://doi.org/10.1080/13504851.2019.1593927. URL: https://doi.org/10.1080/13504851.2019.1593927.

[10] Maarten Goos, Alan Manning, and Anna Salomons. "Explaining job polarization: the roles of technology, offshoring and institutions". In: *Offshoring and Institutions (December 1, 2011)* (2011).

[11] Marc-Andreas Muendler et al. "Job Concordances for Brazil: Mapping the Classificaçao Brasileira de Ocupaçoes (CBO) to the International Standard Classification of Occupations (ISCO-88)". In: *University of California, San Diego, unpublished manuscript* (2004).

[12] Maxim Poletaev and Chris Robinson. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000". In: *Journal of Labor Economics* 26.3 (2008), pp. 387–420. URL: https://EconPapers.repec.org/RePEc:ucp:jlabec:v:26:y:2008:i:3:p:387-420.

A Appendix

A.1 Autonomy

As routine occupations are phased out at higher education levels, the jobs taking their place rely on a different set of abilities, such as autonomy. While routine occupations are by definition inflexible, newer jobs trend towards creativity, problem solving, and communication. Deming (2021) finds that decision-making abilities in particular are increasingly sought after, and lead to abnormally high wage gains, which grow even further throughout each worker's career.

In the QBQ, I use the attitudes dataset and employ the importance score (between 1 and 5, zero when missing) of the attitude "autonomy under work contexts where changes are predicted" as a measure of autonomy. Unlike the routine scores, Figure 15 shows that autonomy is positively associated to qualification level, and that, while nearly absent at the lowest skill levels, a considerable portion of the top two qualification levels are highly autonomous. It also seems to reveal some heterogeneity within education levels: while only 17% of all college-educated workers are in highly autonomous positions, over 35% of workers in the top two qualification levels work in such occupations.

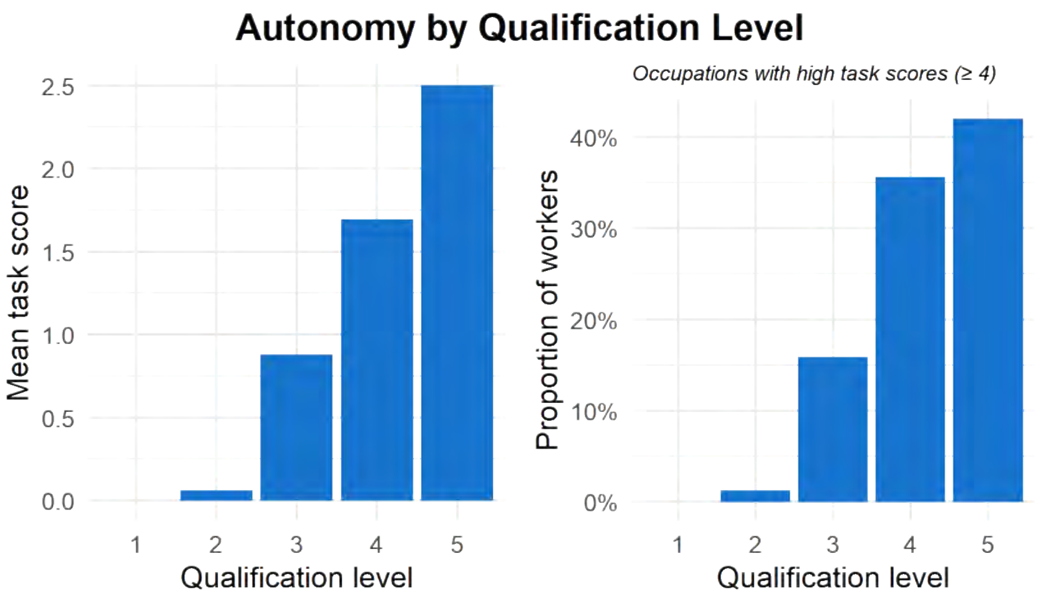


Figure 15

Autonomy is mostly prevalent at professional and also for production occupations, as shown in Figure 16. Since professional occupations grew dramatically, this hints at the conclusion that autonomy should have also had a simultaneous increase. There was no clear pattern of growth in the mean autonomy level.

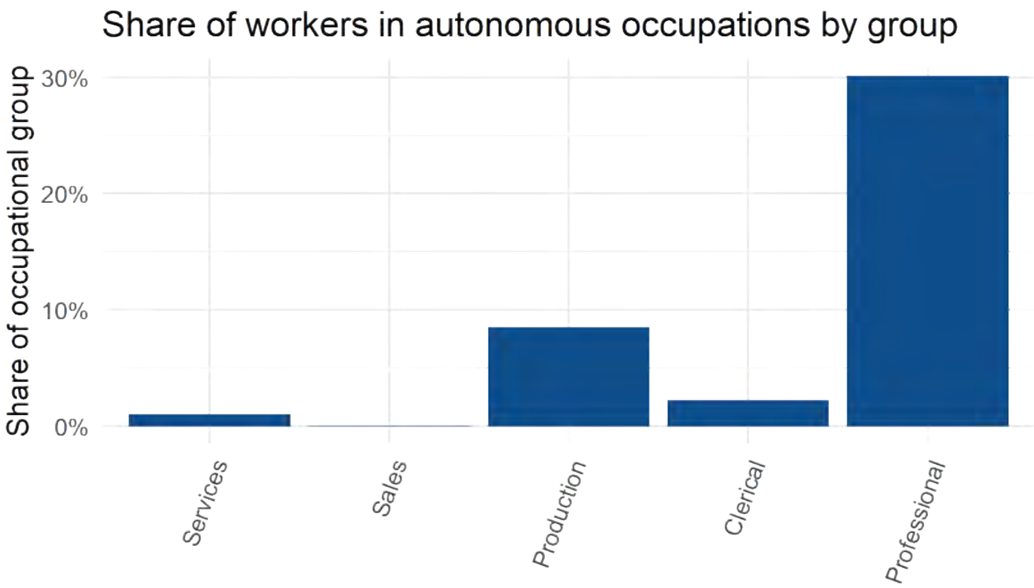


Figure 16

A.2 Abstract Reasoning

Likewise, the level of abstract reasoning used in an occupation can be extracted from the skill "abstract reasoning for problem solving". Abstract tasks are considered by Acemoglu and Autor (2011) to lie at the polar opposite of routine tasks in the distribution of skill between occupations. Under the name of "non-routine analytic", these types of tasks were also deemed to benefit the most from technical change, according to Autor, Levy, and Murnane (2003).

The distribution of this skill up to the qualification levels so far ventured by QBQ is, however, very similar to that of autonomy. It is higher for more highly educated workers in higher-skill occupations, but concentrated among professional and production jobs, with the addition of also being present in some service occupations.

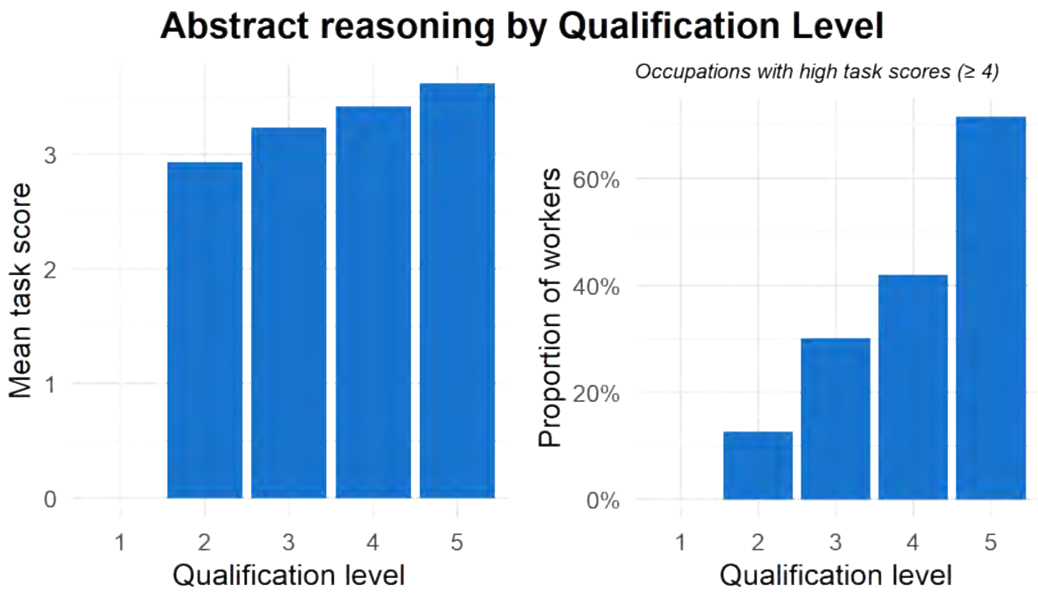


Figure 17

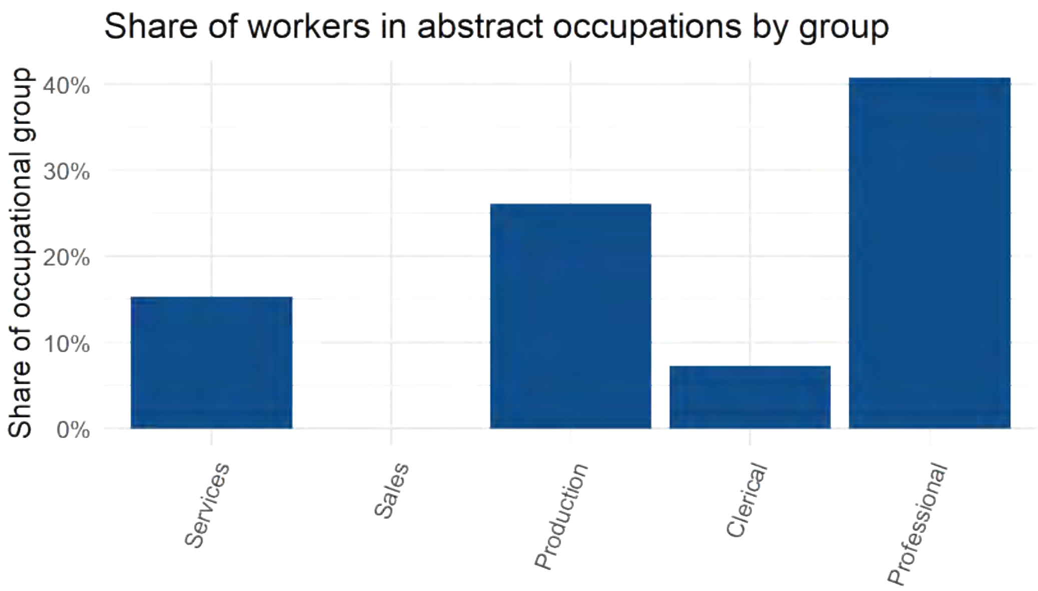


Figure 18

A.3 Interaction

The QBQ has two skill categories relating to interaction: “Communication” and “Interpersonal Skills”. The Interactive score for each occupation was calculated as the mean importance score of skills within either of those categories. This means that, rather than discretely ranging from 0 to 5, this index can take any value in this interval.

Although this measure of interaction is also growing in both skill levels and education, Figure 19 reveals its distribution between occupational groupings to stand out from the others. The ones highest in interaction are sales and professional occupations.

Figure 19

A.4 Technology

The final index, for the use of technology, is drawn from the skill of "applying low-complexity technological principles for problem solving". It, as well as being higher for more skilled occupations, is almost exclusively present in professional and production jobs, as per Figure 20.

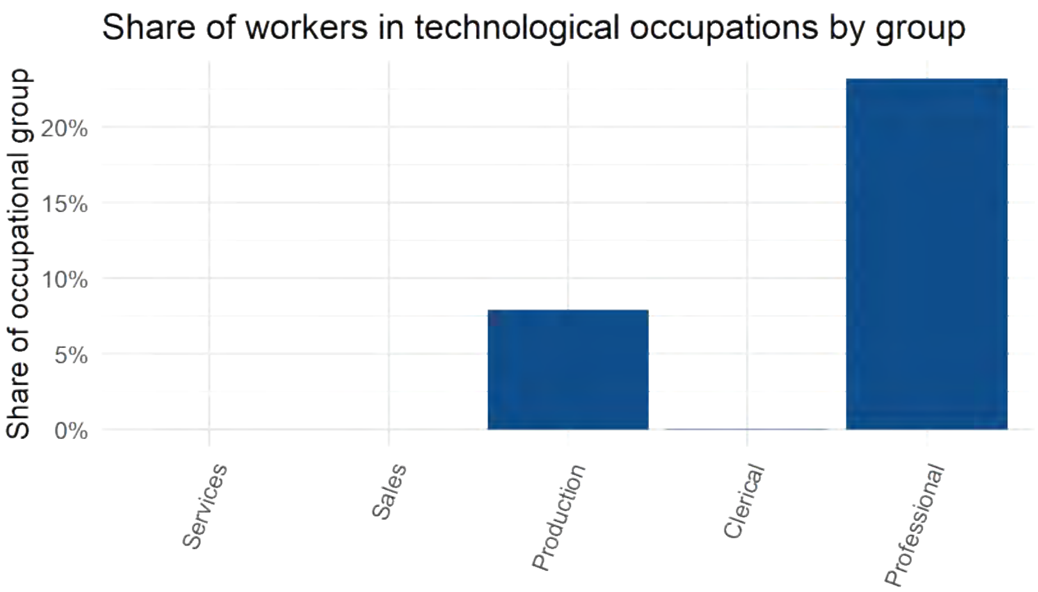


Figure 20