UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

FACULDADE DE CIÊNCIAS ECONÔMICAS

PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

VANESSA NEUMANN SULZBACH

ESSAYS ON JOB POLARIZATION IN THE BRAZILIAN LABOR MARKET

Porto Alegre

2020

VANESSA NEUMANN SULZBACH

ESSAYS ON JOB POLARIZATION IN THE BRAZILIAN LABOR MARKET

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do título de Doutora em Economia.

Orientador: Marcelo de Carvalho Griebeler

Porto Alegre

2020

CIP - Catalogação na Publicação

Sulzbach, Vanessa Neumann
Essays on job polarization in the Brazilian labor market / Vanessa Neumann Sulzbach. -- 2020.
104 f.
Orientador: Marcelo de Carvalho Griebeler.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, BR-RS, 2020.

1. Job polarization. 2. Task content of occupations. 3. Social skills. 4. Task price. 5. Formal and informal sectors. I. Griebeler, Marcelo de Carvalho, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

VANESSA NEUMANN SULZBACH

ESSAYS ON JOB POLARIZATION IN THE BRAZILIAN LABOR MARKET

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do título de Doutora em Economia.

Aprovada em: Porto Alegre, 26 de junho de 2020.

BANCA EXAMINADORA:

Prof. Dr. Marcelo de Carvalho Griebeler - Orientador
UFRGS

Prof. Dr. Adelar Fochezatto
PUCRS

Prof. Dra. Felipe Garcia Ribeiro
UFPEL

Profa. Dra. Thais Waideman Niquito
UDESC

Ao Pedro e a nossa amada Lívia, com amor

ACKNOWLEDGEMENTS

I am very grateful to the Federal University of Rio Grande do Sul and to the professors of the graduate program in Economics for all the teachings and experience sharing. I wish to express my sincere appreciation to my advisor, Marcelo Griebeler, for the continuous support of my doctoral study and research, for his patience, motivation, enthusiasm, and knowledge.

I wish to thank my friends and colleagues (too many to list here but you know who you are!) for providing support and friendship that I needed in this thesis preparation period, especially Guilherme Stein, to who I will forever be thankful for the contributions and discussions of a significant part of the thesis.

I wish to acknowledge the support and great love of my family, my parents Pedro e Maristela Sulzbach, who are my greatest examples; and my brother and sister in law, Bruno Neumann Sulzbach and Isadora Kreuzner de Barcellos. They kept me going on this work with serenity and trust of being suported.

I pay my special thanks to the love of my life, my husband, Pedro Bazani Abrahão, for his support, assistance, advice, companionship and all the affection dedicated to me in every step of this thesis and all of my life. At the most, I dedicate this work to our beloved Lívia, who accompanied me in the final stretch and will soon be with us.

RESUMO

Esta tese apresenta três ensaios sobre polarização dos empregos e dos salários no Brasil. Cada ensaio aborda um tema diferente neste contexto. O primeiro ensaio visa analisar porque a crescente procura por trabalhadores altamente qualificados não tem sido igualmente distribuída entre os gêneros no Brasil. Investigamos a importância das competências sociais na crescente probabilidade de mulheres trabalharem em um emprego bom (*good job*) – aqueles que pagam salários mais elevados – no mercado de trabalho altamente qualificado. Os resultados mostram uma relação positiva entre as competências sociais e a proporção feminina nas ocupações, evidenciando uma escolha natural das mulheres para profissões intensivas neste tipo de competências. Também fornecemos resultados consistentes com a literatura neurocientífica e de psicologia de que as mulheres têm uma vantagem comparativa na execução de tarefas que requerem competências sociais. A relevância dessas competências na determinação dos salários se mostrou mais elevada para as mulheres do que para os homens nos últimos anos. O segundo ensaio propõe a resolução do *puzzle* relativo à polarização dos salários no Brasil. A literatura tem documentado uma tendência descendente dos ganhos salariais ao longo da distribuição de salários, o que contradiz a evidência internacional do padrão de forma em U - a expansão dos salários das profissões de alta e baixa qualificação às custas dos salários das profissões de média qualificação – como resposta ao aumento/diminuição da procura de tais empregos. O ensaio propõe uma forma alternativa de analisar a polarização nos salários através da estimativa dos preços das tarefas (*task price*) utilizando dados de painel a nível individual e utilizando os índices de tarefas contínuos de cada ocupação. Os resultados mostram um aumento acentuado no retorno das tarefas cognitivas entre os dois períodos analisados (2002-2003-2004 e 2012-2013-2014), enquanto que o avanço do preço das tarefas manuais foi mais suave no mesmo período, e o retorno das tarefas de rotina não mudou. Essas evidências sugerem a existência de polarização na estrutura salarial no Brasil, tal como observado em vários outros países. O terceiro ensaio visa pôr luz sobre a tendência divergente nas mudanças de emprego ao longo da distribuição de salários entre os setores formal e informal, propondo uma discussão sobre o impacto do conteúdo das tarefas exigidas por cada ocupação na probabilidade de um indivíduo ser informal e na diferença salarial entre estes dois setores. Os resultados mostram que, mesmo depois de controlar o viés de seleção e por características observáveis dos indivíduos, a probabilidade de ser informal está positivamente correlacionada apenas com a exigências de tarefas manuais. Os requerimentos de tarefas cognitivas e rotineiras estão negativamente correlacionados com a probabilidade de um indivíduo ser informal. Além disso, cognitivo é a tarefa mais importante para explicar a diferença salarial entre os setores formal e informal, tanto em 2003 como em 2015. Sua importância cresceu entre o período, contribuindo para a manutenção do fosso salarial ao longo do tempo. Por outro lado, as tarefas manuais e de rotina têm pouco efeito sobre essa diferença, de modo que podemos concluir que as tarefas mostram um baixo poder de influenciar o fechamento do *gap* salarial formal-informal.

Palavras-chave: Polarização de emprego e salários no Brasil. Conteúdo de tarefas das ocupações. Competências sociais. *Task price*. Setores formal e informal.

ABSTRACT

This thesis presents three essays on employment and wage polarization in Brazil. Each essay addresses a different theme in this context. The first essay aims to analyze why the growing demand for high-skilled workers has not been equally across gender in Brazil. It investigates the importance of social skills in the registered growing probability of women working in a good job – those which pay higher wages – in the high-skilled labor market. The results show a positive relationship between social skills and the female share of occupations, evidencing a natural choice of women for professions intensive in this kind of skill. We also provide results consistent with neuroscience literature that women have a comparative advantage in performing tasks that require social skills. The relevance of such skills in determining wages is higher for women than men in recent years. The second essay proposes to solve a puzzle concerning the polarization of wages in Brazil. Literature has documented a downward trend of wage change throughout Brazil's distribution of earnings, which contradicts the international evidence of U-shape pattern – the expansion of high- and low-skill occupation's wages at the expense of middle-skill ones, in response to increased/decreased demand for such jobs. The essay proposes an alternative way to analyze the polarization in wages by estimating task prices using panel data at the individual level and innovating using each occupation's continuous indexes. The results show a marked increase in the return of cognitive tasks between the two periods analyzed (2002-2003-2004 and 2012-2013-2014), while the advance of manual task price was milder in the same period, and the return of routine tasks did not change. The findings suggest polarization in the wage structure in Brazil, as observed in several other countries. The third essay aims to put light on the divergence trend in employment changes through the earnings distribution between formal and informal sectors, by proposing a discussion about the impact of occupational task content on the probability o an individual being informal and n the wage gap between these two sectors. The results show that, even after controlling for selection bias and observables characteristics, the probability of being informal is negatively correlated to cognitive and routine while the only positive effect among the three tasks comes from manual ones. Also, cognitive is the most important task to explain the gap both in 2003 and 2015, and its importance grew between the period, contributing to the maintenance of the wage gap over time. On the other hand, manual and routine tasks have little effect on it, so that we can conclude that tasks show a low power to influence the closing formal-informal wage gap.

Keywords: Job polarization in Brazil. Task content of occupations. Social skills. Task price. Formal-informal sectors.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

In this thesis, we analyze the polarization of the Brazilian labor market. Over the past two decades, expanding literature has focused on understanding how technological changes have affected the demand for jobs. The growing demand for both the higher skill levels jobs and the low-skilled ones, in contrast with the relative disappearing of middle-class jobs (which requires a moderate level of skills), has been attributed to the rapid adoption of computers.

The seminal study by Autor, Levy and Murname (2003) showed the dichotomous analysis between low and high ability (whose proxy was years of schooling), which prevailed in the previous literature, no longer supported the change in the dynamics of jobs. It was necessary to look at the tasks required for each job since the rapid reduction in the price of computers caused them to replace jobs that demanded routine tasks (easy to code) and began to increase the demand for cognitive (complementary tasks) and manual tasks (difficult to replace). The final equilibrium movement has been called the polarization process because of each task's position in the wage distribution — cognitive in the upper tail, manual in the bottom, and routine in the middle. The vast literature has followed the ideas of Autor, Levy and Murname (2003), most of all, engaged in understanding the phenomenon in developed countries.

In emerging markets, there are fewer studies concerning the subject. A complete analysis was made by Reijnders and Vries (2017), which documented that labor markets also polarized in most emerging countries studied. Although there is no consensus on the phenomenon in Brazil, the Brazilian labor market data indicates that there has been, in recent years, an increase in demand for workers who perform cognitive and manual tasks to the detriment of routine tasks. We explored this evidence throughout the thesis.

Some recent works have explored the fast-expanding of the tasks set performed by computers, substituting even some cognitive tasks, making human interactions (or social skills) more valued. In the first essay, we show that the growing demand for high-skilled workers has not been equally across gender in Brazil. We explore the hypothesis that women have been favored by the increase in demand for social skills in this market. The hypothesis is based on the psychology and neuroscience literature that women have comparative advantages when performing social skills.

We demonstrate a positive relationship between social skills and the female share of occupations, evidencing a natural choice of women for professions intensive in this kind of skill. We also provide results consistent with neuroscience literature that women have a comparative advantage in performing tasks that require social skills. The relevance of such skills in determining wages is higher for women than men in recent years.

Job polarization literature also provides evidence of the simultaneous expansion of high- and low-skill occupation's median wage at the expense of middle-skill occupation's wages, in response to increased/decreased demand for such jobs. So, the U-shape pattern has also been documented in wages structure. Brazilian literature, in contrast, documented a downward trend of wage change throughout Brazil's distribution of earnings, which contradicts the international evidence. A hypothesis that may explain this trend is that employment shifts might result in a temporary imbalance on the worker's average characteristics of each occupation, such that using occupational mean (or median) wages may not be the best way to assess the polarization phenomenon on the wage distribution.

In the second essay, we propose an alternative way to analyze the polarization in wages by estimating task prices using panel data at the individual level to extract the exact price of each task. This paper's innovation is to take into account the continuous indexes of each occupation instead of segregating them in broad exclusive groups as the massive literature has done. By doing so, we preserve information on occupational task requirements, especially in occupations with a high level of two of the three tasks. With all the information available, the task price estimate becomes more accurate.

We show a marked increase in the return of cognitive tasks between the two periods analyzed (2002-2003-2004 and 2012-2013-2014), while the advance of manual task price was milder in the same period, and the return of routine tasks did not change. The findings suggest indications of polarization in the wage structure in Brazil, as observed in several other countries.

Theoretical basis and the international empirical literature of job polarization do not devote much attention to the consequences of the migration from informality to formality. In Essay III, we propose a discussion about the occupational task content and the divergence between formal and informal sectors. We present an estimation of the impact of cognitive, routine, and manual tasks in the probability of an individual being informal and in the wage gap between these two sectors.

We perform a three-stage model for estimating the role of tasks on the likelihood of being informal using PNAD data. We first perform a Heckman (1979) two stages to simultaneous determine earnings and selection between formal and informal sectors and then estimate structural probit. The results show that even after controlling for selection bias and observables characteristics, the probability of being informal is negatively correlated to cognitive and routine while the only positive effect among the three tasks comes from manual ones.

The results presented in this thesis prompt several threads of future research. First, the evidence about the positive correlation between social skills and the female share of occupations – which indicates a natural choice of women for professions intensive in this kind of skill – and the importance of such skills on women's wages open a promising research pipeline. Secondly, the task price approach also provides a new path for research on wage polarization in Brazil, solving the puzzle so far unsolved by literature. Furthermore, the evidence of tight relation between manual tasks and informality jobs (even after controlling for selection bias and observable characteristics) provides subsidies for future analyses on the hypothesis that increase in employment in the lower tail of the wage distribution of the formal sector of the Brazilian labor market may have come from the formalization process in Brazil.

2 ARE SOCIAL SKILLS HELPING WOMEN IN THE HIGH-SKILLED BRAZILIAN LABOR MARKET?

Abstract

In this paper we analyze why the growing demand for high-skilled workers has not been equally across gender in Brazil. It investigates the importance of social skills in the registered growing probability of women working in a good job – those which pay higher wages – in the high-skilled labor market. The results show a positive relationship between social skills and the female share of occupations, evidencing a natural choice of women for professions intensive in this kind of skill. We also provide results consistent with neuroscience literature that women have a comparative advantage in performing tasks that require social skills. The relevance of such skills in determining wages is higher for women than men in recent years.

2.1 INTRODUCTION

The job market has changed rapidly in a few past years in response to technological changes. A vast international literature has documented a rising demand for high-skilled labor (who are apt to perform cognitive tasks) and a fall in demand for workers performing routine tasks. The main findings compose the *job polarization* literature, which indicates a rise in demand for manual tasks as well. However, few studies are looking at gender differences in this context.

In this paper, we show that the growing demand for high-skilled workers has not been equally across gender in Brazil. The likelihood of college-degree women working in a *good job* increased between 1994 and 2017, while the probability for men has fallen. By *good jobs*, we mean the occupations with a median wage in the top twenty percentile wage ranking in the previous year. The massive incorporation of women population in the labor market is well known. Everything else constant, wages are weakly decreasing in response to a rise in the female labor supply. However, this evidence shows that the number of women earning high wages increased, which leads us to believe that there was an increase in demand for women in the Brazilian labor market.

With this evidence in mind, we first investigate the role of the increase in women's labor supply and the possible shifting in demographic and spacial characteristics of the individuals in the gender different trends in the high-skill labor market. Our findings indicate that there are other important things behind this phenomenon beyond the rise in the female share of the college-educated population and the compositional changes in demographics and regions.

To explore the potential drivers of the increase in the likelihood of women choosing *good jobs*, we follow the theoretical model presented by Cortes, Jaimovich and Siu (2018). The model suggests two channels for the phenomenon. The first one refers to the falling discrimination against women in the high-skill labor market. The second is about the rising demand for specific female labor in this market (what they call *Female Bias*). The hypothesis is that women have been more demanded because they are best in performing social skills since they have a comparative advantage in jobs that require them. Also, the returns of these tasks have increased recently in the labor market, which would explain why the probability of women working in better paid jobs increased.

Psychology and neuroscience literature indicates a natural ability of women performing social tasks. There are sex developmental differences in brain structure and function that lead women to have more empathy and be better in interacting with others (CHAPMAN et al., 2010). The ability to understand what others are thinking and feeling enable women to be better constructing work teams and *collective intelligence* (WOOLLEY et al., 2010).

Prominent literature has been exploring the growing importance of social skills in the labor market. In this context, social skills are associated with leadership, communication, and interpersonal skills. In the workplace, the ability to interact with other people, and understanding the overall environment has become more valuable because it increases team productivity. Several studies have presented evidence of the increasing demand for social skills by showing theses skills developed in high school are a stronger predictor of employment and wages for young adults. Besides that, recent works have highlighted the social skill valorization in the labor market as a result of the difficulty of computers on replacing human interactions, like "reading minds of others and reacting" (DEMING, 2017). Also, it has been related to the closing gender-wage gap (BORGHANS; WEEL; WEINBERG, 2014).

We investigate the hypothesis of *Female Bias* in the Brazilian labor market, given that the first channel of the model (declining of female discrimination) is quite difficult to measure. The central assumption of our work is that the appreciation of social skills has benefited women in the Brazilian labor market because they have a comparative advantage in this. We contributed to the literature by building a novel Brazilian Classification of Occupations (CBO[1]) rating based on skill content compatible with the O*NET[2] task index. We follow the assumption that task content is similar across countries for similar occupations. We match the skill intensity content of occupation with Annual Report

[1] Classificação Brasileira de Ocupações - CBO

[2] US Department of Labor's Occupational Information Network database (O*NET)

on Social Information (RAIS[3]) database to investigate whether women's comparative advantage hypothesis holds in the Brazilian labor market. To do so, we evaluate whether: (i) women naturally choose intensive social occupations; and (ii) the returns of social skills are higher for women than men.

Our results show a positive relationship between social skills and the female share of occupations. Because such a share is negatively correlated to other tasks such as cognitive, routine, and manual, our findings suggest the possibility of women's natural selection into activities intensive in social skills. We also provide results consistent with neuroscience literature that women have a comparative advantage in performing tasks that require social skills. The relevance of such skills in determining wages is higher for women than men. Social skills also raise cognitive returns in both cases, having increased its effect between 1994 e 2017.

Besides this introduction, this paper is organized into six other sections. In the next one, we discuss the literature on the growing importance of social skills and the female comparative advantage. In the third section, we show evidence of the different gender trends in the high-skilled labor market. Section four describes the model on which the main idea of this paper is based on and presents its predictions. In section five, we present the data used in this paper, detail the construction of the novel CBO rating based on skill content compatible with the O*NET, and show some descriptive statistics of the Brazilian high-skilled labor market. We present the empirical strategy to investigate the model predictions in Brazil and its results in section six. Section seven concludes.

2.2 RELATED LITERATURE

The importance of communication skills in the labor market started to be studied with the work of Autor, Levy and Murname (2003). According to the authors, computer adoption increased the labor productivity of workers engaged with communication tasks (direction, planning, and control) and problem-solving tasks (mathematics). They were the first separating tasks from skills in a "simple economic model used to predict how demand for workplace tasks responds to an economy-wide decline in the price of computer capital." Until their findings, the literature separated high-skill workers (with at least college education) and low-skill workers (others) and looked only at the impact of technological change on these two groups (supposing a one-to-one map between tasks and skills).

The main findings of their study suggest an increase in employment and earnings of low and high ability workers who perform cognitive and manual tasks (meaning the occupations on the lowest and higher tail of the wage distribution) to the detriment of the medium ability worker, who perform routine tasks (or the occupations that are in

[3]Relação Anual de Informações Sociais - RAIS

the middle of the wage distribution). That is, the relation between demand for skills and the wage distribution was no longer monotonic and became to be in a new format: U shape. This phenomenon is known as the polarization of employment and wages. Numerous researchers have documented it (Autor, Katz and Kearney (2006); Spitz-Oener (2006); Goos and Manning (2007); Acemoglu and Autor (2011); Autor and Dorn (2013); Goos, Manning and Salomons (2014); Ikenaga and Kambayashi (2016)).

As a new production factor, with a rapidly declining price, the computer/machine performs tasks with well defined and explicit rules, replacing workers allocated in jobs that are intensive in routine tasks, as clerical, administrative, production, and operative occupations (ACEMOGLU; AUTOR, 2011, p. 35).

On the other hand, computers have specific and limited capabilities. Their ability to perform tasks depends on workers with sufficient capacity to determine which decision rules will be valid in computational processes. The tasks performed by these workers are cognitive. In this case, they involve complex problem solving and require a certain level of communication, which computers are unable to perform. Besides, computers have a limitation on performing manual tasks that require physical dexterity. Even with the advance of the technology developing robots to do some manual tasks (with more precision than human), human labor performing these tasks is still a cheap production factor to be replaced for those robots. So, it seems that there is a complementary relation between computers and cognitive/manual tasks. Job polarization literature argues the de-skilling process, in which routine workers moved down the occupational ladder and have begun to perform jobs traditionally performed by lower-skilled workers in manual tasks.

New literature explores the fast-expanding of the tasks set performed by computers. Over time, they began to replace some cognitive tasks too. The idea is that in the implementation of IT Revolution, cognitive tasks were a key component of the capital investment phase once they were essential to build and maintain the new capital. So its demand increased rapidly. However, with the maturity of this investment, the demand for cognitive tasks begin to narrow with computer also substituting some human interactions (BEAUDRY; GREEN; SAND, 2016). In the recent past, automation using explicit rules or manually written computer algorithms were limited to areas where there was explicit knowledge to codify. Now, machine learning (ML) and artificial intelligence (AI) are substituting even high-skill labor. “Recent rapid progress in ML has been largely driven by an approach called deep learning and has made it possible for machines to match or surpass humans in certain types of tasks, especially those involving image and speech recognition, natural language processing, and predictive analytics”(BRYNJOLFSSON; MITCHELL; ROCK, 2018).

Deming (2017) suggest there has been a reduction in science, technology, engineering, and mathematics (STEM) employment and earnings nos EUA and growing importance of the other cognitive occupations – such as managers, nurses, teachers, lawyers, economists, and others – which require interpersonal interaction. In this context, social skills have gained relevance once some human interactions are still difficult to automate, taking some relative advantage against other skills in the labor market (AUTOR, 2015). The ability to listen and read and instantly react or make a decision are still challenges for computers. More than that, the ability to "read between the lines" and to feel the best time to interfere or make a decision are exclusively human characteristics (so far).

Besides the advantage of social skills for being difficult to be replaced by computers, it has gained value in the market for increasing the productivity of firms. According to Deming (2017), team productivity increases when some people have social ability, once it reduces coordination costs and allows workers to specialize in their bests tasks, increasing overall productivity. For the United States labor market, the authors show that high-paying jobs have increasingly required social skills because of their capacity to help raise the productivity of other factors, including cognitive labor.

Several works have been exploring the growing importance of the ability to interact with or handle interactions with people. Most of them explore adolescent skills endowment from high school surveys and match with adult labor market outcomes. With this methodology, Borghans, Weel and Weinberg (2014) report employment and earnings grew in occupations requiring "people skills" in Britain, Germany, and the United States. Others construct models to capture social skills returns in the labor market. In the model of Kambourov, Siow and Turner (2014), workers differ from each other by the level of "relationship skills". Those with higher levels are associated with higher earnings and a stable marriage. In the multi-sector model of McCann et al. (2010), there are productivity gains from specialization and team production, but it requires communication and coordination between team members, which makes social skill appreciate once it rises the productivity of cognitive workers. The complementarity between cognitive and social skills is also highlighted by Weinberger (2014), who also link the pre-labor skills to adult outcomes.

The rise in demand for social skills has also been related to the closing on the gender wage gap, in the sense of increasing earnings in occupations that require social skills have benefit women more intensively relative to men (BACOLOD; BLUM, 2010; BORGHANS; WEEL; WEINBERG, 2014; BLACK; SPITZ-ONER, 2010). The fundamental hypothesis is grounded on the psychology and neuroscience literature, which has proved that women are better performing interactions with other people and care more about others (GILLIGAN, 1993). Authors such as Woolley et al. (2010) argue that the presence of women in teams/groups increases what they call *collective intelligence*, which is not affected by its member's cognitive level (individual intelligence), but by the social sensitivity of the group. The reports also indicate the female advantage in decoding nonverbal communication

(HALL, 1978).

Other authors argue that the sex difference in social abilities is related to the intrauterine development of brain structure and function. For Chapman et al. (2010), females have, on average, stronger drive to empathize than males, which involves a better understanding of what others are thinking and feeling, and, therefore, a better aptitude for interacting in the social world. It is defined, according to authors, by the level of fetal testosterone. In the same line, Baron-Coen, Knickmeyer and Belmonte (2005) shows that women are stronger in empathy while men are better systemizing; that is, women are better in predicting and responding to agents' behavior.

In the context of analyzing gender-specific effects of technological changes, Cortes, Jaimovich and Siu (2018) find that women may have benefited from the increased demand for social skills because of their comparative advantage in these skills, as advocated by psychology and neuroscience literature. They present a simple equilibrium model of the high-skilled labor market, considering the traditional neoclassical framework of labor demand and supply. They separate labor inputs into *good jobs* and others and rationalize the firm's problem considering discrimination towards women in the labor market. Since workers differ in the ability to work in a *good job*, the clearing of the market depends on that heterogeneity. From the model's equilibrium, it is possible to decompose the gender difference trend in the probability of working in *good job* into (i) a component related to the discrimination of the "two markets" (*good jobs* and other jobs); and (ii) another factor related to the demand for each gender in both markets. The details of the model and its predictions are present in Section 2.4.

Cortes, Jaimovich and Siu (2018) also perform an interesting empirical analysis. The authors use estimations of labor demand for social skills obtained by the specific attributes required in job advertisements (which database is made available by Atalay et al. (2018)) to construct "the change in the importance of social skills of each occupation (i.e., change in demand for social skill)". They find a positive relation with the increase in female share, which corroborates their theoretical model's predictions.

Given the absence of such information on job requirements and of adolescent skills endowment survey for Brazil, we take as given the evidence of social skills valorization and increasing demand for them in the Brazilian labor market. It is plausible to make such an assumption once the process of replacing some jobs with machines could take little different time to a country to another, but in more extended periods, such as 20 years, the process tends to be similar. Nevertheless, it is true that institutions matter, such that the effects of the technological changes on developing countries are not identical to developed ones'. It is not possible yet to substitute bus ticket collectors – an occupation that has disappeared in other countries – in some municipalities in Brazil, for example. However, most job replacements by machines have been widespread in the last two decades as bank tellers, online sales (began with airplane ticket sales and spread to e-commerce), operating systems in enterprises, etc.

Despite of the scarcity of literature on skills in the Brazilian labor markets, we can find studies that reinforce that our assumption about the behavior of social skills. For instance, Adamczyk, Franca and Fochezatto (2019) endorsed the idea of the difficult substitutability of social skills by computers by evaluating the Brazilian occupations accordingly to its automation probabilities. The authros found that "tasks that require high levels of social and creative intelligence are seen as less likely to be substituted by machines", so that such an assumption appears reasonable.

Finally, as the evidence of female comparative advantage in social skills is well documented by psychology and neuroscience literature, we assume this is true for Brazil. By taking these two assumptions as true, we test the hypothesis of *Female Bias*, as predicted by the model of Cortes, Jaimovich and Siu (2018), in the Brazilian labor market.

2.3 HIGH-SKILLED LABOR MARKET: SOME EVIDENCES

In this section, we present some evidence that movements in the high-skilled labor market have not been equal across genders. In particular, we show that there has been an increase in women's employment in *good jobs* and a fall in the likelihood that a college-educated male is employed in a high-wage/cognitive occupation. The Brazilian labor market is characterized by a clear separation between high and low skilled driven occupations, which implies that some *good jobs* are only available for more educated workers. For this reason, we focus on the high-skilled labor market.

We consider high-skilled workers those who have at least a college degree in terms of educational attainment. *Good job* are defined as occupations whose median wage is in the highest decile of the 1994 occupational wage distribution. Both definitions follow the ones adopted by Cortes, Jaimovich and Siu (2020). Our analysis uses microdata from the RAIS database for the years 1994 to 2017, provided by the Ministry of Economy – Special Secretary of Labor. Data is detailed in Section 2.5.

The high-skilled population increased by nearly 2 million between 1994 and 2017, more than half of women (1.2 million). Despite the substantial increase in the high-skilled labor supply, the probability that a high-skilled worker was employed in a high-paying occupation (*good job*) dropped over the same period from 43.9% to 39.9%. This decrease, however, hides divergent trends across genders. Table 1 presents high-skilled occupational employment and shows the key statistics motivating our analysis. In 1994, the probability of a woman working in a *good job* was 24.5% and expanded to 33.2% over the following 23 years until 2017. By contrast, the portion of high-skilled men working in *good jobs* was about 56.1% in 1994 and has fallen to 47.1% since that. This evidence is particularly striking given the massive increase of women population (women's labor supply) in the high-skilled labor market.

The divergent gender trends in the probability of working in *good jobs* imply that there is an increase in the female share of employment in these jobs. Figure 1 present some pieces of evidence of such movements. Each circle represents a 4-digit occupation (CBO 2002) and its size indicates the occupation's share of aggregate employment in 1994. Along the horizontal axis, occupations are ranked by their place in the 1994 wage per hour distribution[4]. On the vertical axis is the 1994-2017 change in the female share of high-skilled employment[5] by occupation.

Table 1 – High-skilled occupational and employment status - 1994-2017. Top 20 wage percentile as *good job*

High-Skilled Workers		1994	2017	Difference Total	Difference %
Total	Number	681,912	2,897,600	2,215,688	324.9
	Top 20% (%)	43.9	39.9	-4.0	-9.1
	Others (%)	56.1	60.1	4.0	7.1
Women	Number	264,013	1,508,231	1,244,218	471.3
	Top 20% (%)	24.5	33.2	8.7	35.7
	Others (%)	75.5	66.8	-8.7	-11.6
Men	Number	417,899	1,389,369	971,470	232.5
	Top 20% (%)	56.1	47.1	-9.0	-16.0
	Others (%)	43.9	52.9	9.0	20.5

Source: Elaborated by the author from RAIS data, 18-65 years old employees from private sector with at least college degree. Employment categorized by ranking in occupational wage distribution of 1994.

Figure 1 shows the increased relative sorting of women to *good jobs*. Although the proportion of women has increased in all occupations, higher-paying ones experienced a bit larger in the female proportion of employment between 1994-2017. The positive association has an estimated coefficient that is significant at the 1% level. The results suggest the changing occupational choices of men and women may have played an important role in explaining the increase in female share employment in *good jobs*. The rise in the college-educated female population does not seem to be the only answer.

Given that evidence, we are interested in analyzing whether this divergence is due to (i) the demographic characteristics of men and women in the high-skilled labor market; or (ii) a compositional shifting across occupations in the high-skilled market between men and women. To address the first investigation we perform an Oaxaca-Blinder Decomposition separately per gender to verify whether the change in probability of working in a *good job* can be attributed to changes in the demography and spacial characteristics. To shed

[4]Occupation's wage per hour is calculated by the median wage per hour of the workers from each 4-digit occupation.

[5]For visual clarity, Figure 1 excludes seven occupations where the change in the female share of high-skilled employment increased more than 40 percentage points and five occupations where the change fell more than 40 percentage points. The fitted regression line is based in all occupations, and its coefficient is significant at the 1% level

Figure 1 – Change in female employment share among colleges graduates and occupational wage ranking

Source: Elaborated by the author from RAIS data. Notes: The figure plots log changes in employment shares and occupational median wage per hour by 1980 occupational skill percentile rank using a locally weighted smoothing regression. Skill percentiles are measured as the employment-weighted percentile rank of an occupation's median log wage in RAIS dataset. Consistent occupation codes (CBO codes) for 1994 and 2017 were used using method decribed in subsection 2.5.1 of this thesis's essay I.

light on the second possibility, we decompose the rising movement of the female share of high-skilled employment into a component exclusive to the rise of college-educated women population and another component related to the occupational choice of men and women.

2.3.1 Decomposition exercises

In this section, we present two decomposition exercises. The first one investigates the hypothesis of the trend difference in the probability of working in a good job is related to the demographic characteristics of men and women in the high-skilled labor market. We do so by running the Oaxaca-Blinder decomposition technique. In the second one, we present a decomposition of a rising movement of the female share of high-skilled employment into two components: (i) the rise of college-educated women population; and (ii) occupational choice of men and women.

2.3.1.1 Oaxaca-Blinder Decomposition

In this subsection we adapt the Oaxaca-Blinder decomposition technique (OAXACA, 1973; BLINDER, 1973) to the framework of our interest in investigating some evidences of high-skilled market. We process the decomposition for all individuals, for men, and women. The probability of working in a *good job* is observed in two periods: 1994 ($s_i = 0$) or 2017

$(s_i = 1)$. It is a function of individuals characteristics, which formally can be written as:

$$\begin{aligned} \pi_i^s &= X_i\beta_1 + \varepsilon_{1i}, \text{ if } s_i = 1 \\ \pi_i^s &= X_i\beta_0 + \varepsilon_{0i}, \text{ if } s_i = 0 \end{aligned} \tag{2.1}$$

where π_i^s represents the probability of working in a *good job* in year $s \in \{1, 0\}$ and X_i is the vector composed of individual characteristics. Moreover, we have that $E[\varepsilon_{1i}|X_i, si] = E[\varepsilon_{0i}|X_i, si] = 0$.

The difference in the probability between 1994 and 2017 is given by:

$$E[\pi_i^s|si = 1] - E[\pi_i^s|si = 0] = (E[Xi|si = 1] - E[Xi|si = 0])\beta_0 + E[Xi|si = 1](\beta_1 - \beta_0) \tag{2.2}$$

As one can see in the Table 2, demography and spacial characteristics (*explained component*) do not account for a large part of the difference in the probability of a high-skilled employee (total) working in a *good job*. Variables such as age[6] and region of residence contribute only with 9.8% to the decrease observed between 1994 and 2017. We see a similar pattern for men (7.1%). In the case of women, for whom the probability difference increased, the explained part has even less importance (5.4%). Aging helped high-skilled women to increase the probability of working in a *good job* (5.2%), while this variable contributed to reducing the gap for the case of men and total.

Our results indicate that compositional changes in demographics and regions are insufficient to explain the divergent trend in probability of working in *good jobs*, which seems to be due predominantly to changes in the propensity to work in them, conditional on observable characteristics.

2.3.1.2 Decomposing the rise movement of women in high-skilled labor market

The massive increase of women college-educated population over the past years seems to be the most critical driver of the rising female share of high-skilled employment in a *good job*. However, is that the only reason for the increasing movement? What is its importance on the overall advance of female share in *good job*?

As shown by Cortes, Jaimovich and Siu (2018), the female share of high-skilled employment in *good jobs* can be decomposed as:

$$\sigma_t = \frac{F_t^{good}}{F_t^{good} + M_t^{good}} = \frac{\overline{F_t} * \pi_t^F}{\overline{F_t} * \pi_t^F + \overline{M_t} * \pi_t^M} \tag{2.3}$$

[6] We only use age as a demographic variable because information of race or physical disability are not available before 2003. We also don't use nativity, as Cortes (2016) did, because there were few variation among individuals.

Table 2 – Decomposition of the change in the probability of working in a *good job*

	Total		Women		Men	
	p.p	%	p.p	%	p.p	%
Difference	-3.99 (0.0006)***	100.0	8.73 (0.0009)***	100.0	-9.00 (0.0008)***	100.0
Unexplained	-3.60 (0.0006)***	90.2	8.26 (0.0009)***	94.6	-8.36 (0.0008)***	92.9
Explained	-0.39 (0.0001)***	9.8	0.47 (0.0002)***	5.4	-0.64 (0.0001)***	7.1
age	-0.37 (0.0001)***	9.2	0.45 (0.0001)***	5.2	-0.66 (0.0001)***	7.3
south	-0.10 (0.0000)***	2.5	-0.11 (0.0000)***	-1.2	-0.05 (0.0000)***	0.6
north	0.02 (0.0000)***	-0.6	0.04 (0.0000)***	0.4	0.00 (0.0000)*	0.0
northeast	-0.06 (0.0000)***	1.5	-0.05 (0.0000)***	-0.5	-0.02 (0.0000)***	0.2
midwest	0.12 (0.0000)***	-2.9	0.14 (0.0000)***	1.6	0.09 (0.0000)***	-1.0

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05 (***), p<0.01. We only use age as a demographic variable because information of race or physical disability are not available before 2003. We also don't use nativity, as Cortes (2016) did, because there were few variation among individuals.

where F_t^{good} is number of women in a *good job* (median wage in top 20%) in time t, M_t^{good} is number of men in a *good job* in time t, $\overline{F_t}$ is the total number of high-skilled (college-educated) women, $\pi_t^F \equiv \frac{F_t^{good}}{\overline{F_t}}$ is the fraction of high-skilled women employed in *good job*. By similarity, $\pi_t^M \equiv \frac{M_t^{good}}{\overline{M_t}}$ is the same for men.

Table 3 shows the counterfactual exercises. Among all high-skilled workers, the probability of women working in a *good job* raised from 21.6% to 43.3%. The 21.7 percentage point increase is explained by a component due to high-skilled population growth (mainly women) and another element due to changes between occupations (increasing the proportion of women in top decile wage per hour occupations). The column signed by 2017* considers the effect of the rise in the female high-skilled (college-educated) population. In this case, the counterfactual is built holding π_t^F e π_t^M at levels of 1994, allowing only the number of high-skilled men and women to vary. The increase in the female high-skilled population has an important role. If only this component had happened, the probability of women working in *good jobs* would be 32.1%, explaining about 48.4% of the rise. Thus, the women's effort to be more educated does have impacted the female share of *good job*.

The column signed by 2017**, on the other hand, considers the women's occupational choice; that is, it takes into account the rise of female choice for higher-paying occupations. To construct that, we keep the high-skilled population at the level of 1994, allowing only the fractions π_t^F e π_t^M to vary. If only the women's migration to *good jobs* had happened (without a rise in women population), the proportion of women in these occupations would be 30.8%, accounting for 42.3% of the increase. Im sum, the population and the occupational choice effects have played a similar role.

The results are similar to what Cortes, Jaimovich and Siu (2018) had found to the United States, where the population's contribution was 36.1%. Occupational choice has a bit larger effect there (49.2%). In both cases, the evidence shows there are other important aspects beyond the rise in the female share of the college-educated population. The change in the men's and women's occupational choice equilibrium is something that we have to look at properly.

Table 3 – High-skilled female share of employment and counterfactual exercise

	Observed				Counterfactual	
	1994	2017	Variação p.p	Variação %	2017*	2017**
Population vs Occupational Choice					Population	Occupation
Top 20% (%)	21.6	43.3	21.7	100.6	32.1	30.8
Bottom 90% (%)	52.1	57.8	5.7	11.0	65.1	44.4

Source: Elaborated by the author from RAIS data, 18-65 year old employees from private sector with at least college degree. Notes: The counterfactual 2017* considers the effect of the rise in the female high-skilled (college-educated) population. It is built holding π_t^F e π_t^M at levels of 1994, allowing only the number of high-skilled men and women to vary. The counterfactual 2017** considers the women's occupational choice by higher-paying jobs. It is built keeping the high-skilled population at the level of 1994, allowing only the fractions π_t^F e π_t^M to vary.

2.4 EQUILIBRIUM MODEL OF THE HIGH-SKILLED LABOR MARKET

We follow the equilibrium model of the market for high-skilled workers presented by Cortes, Jaimovich and Siu (2018), which summarizes the predictions of neoclassical theory and allows for gender-specific characteristics. The supply of workers, labor productivity, and discrimination are different between male and female workers.

The output of the economy is produced by a combination of high-skilled (college-educated) labor and other inputs, such as low-skilled labor, capital, etc. The real output is given by:

$$Y_t = \Psi(f^G(Z_{Mt}^G L_{Mt} + Z_{Ft}^G L_{Ft}), f^O(Z_{Mt}^O E_{Mt} + Z_{Ft}^O E_{Ft}), K) \tag{2.4}$$

where L_{gt} is the high-skilled labor working in *good jobs*, with $g = \{M, F\}$ where M stands to male and F to female. By similarity, E_{gt} is the high-skilled labor working in other occupations. The vector K gives other factor inputs, including capital, low-skilled labor, etc. High-skilled individuals are endowed with different abilities in *good jobs*. They will sort into *good jobs* (L_{gt}) and others ones (E_{gt}) according to their abilities.

Effective labor is augmented by productivity. In this case, the productivity of a high-skilled workers is gender-specific so that Z_{Mt}^i stands for male and Z_{Ft}^i for female, with $i = G, O$, where G refers to *good jobs* and O refers to other jobs.

We assume that effective labor inputs of high-skilled men and women are perfect substitutes, so marginal rates of transformation between male and female labor are constant. That is, $f^G(.) = f^G(Z^G_{Mt}L_{Mt} + Z^G_{Ft}L_{Ft})$ and $f^O(.) = f^O(Z^O_{Mt}E_{Mt} + Z^O_{Ft}E_{Ft})$. The function Ψ is assumed to be constant returns to scale, with $\Phi_1 > 0$, $\Phi_2 > 0$, $\Phi_{11} < 0$, $\Phi_{22} < 0$,$f^i_1 > 0$,$f^i_2 > 0$, $f^i_{11} \leq 0$, $f^i_{22} \leq 0$ for $i = G, O$.

Labor Demand: The labor market we are dealing with is competitive. Although the labor inputs of high-skilled men and women are perfect substitutes, there is discrimination against women in this market. As it was first thought in the seminal work of Becker (1974), discrimination is modeled here as tax for the hiring firms. The firm's problem is:

$$\max_{L_{Mt},L_{Ft},E_{Mt},E_{Ft},K_t} Y_t - (1+\tau^G_t)w_{Ft}L_{Ft} - w_{Mt}L_{Mt} - (1+\tau^O_t)p_{Ft}E_{Ft} - p_{Mt}E_{Mt} - \mathbf{r}_tK_t \quad (2.5)$$

where $(1+\tau^G_t)$ is the discriminatory parcel towars high-skilled women in the "good jobs", which may differ from that one in the other occupations $(1+\tau^O_t)$; w_{gt} is the wage earned by a high-skilled worker in the *good jobs* and p_{gt} is the same for the other occupations; $\mathbf{r}$ is the vector of other input's prices.

The first order conditions gives us the labor demand functions for L_{Mt}, L_{Ft}, E_{Mt} and E_{Ft}:

$$\begin{aligned} w_{Mt} &= Z^G_{Mt}\Psi_1(.)f^G_1(Z^G_{Mt}L_{Mt} + Z^G_{Ft}L_{Ft}) \\ w_{Ft} &= \frac{Z^G_{Ft}}{(1+\tau^G_t)}\Psi_1(.)f^G_2(Z^G_{Mt}L_{Mt} + Z^G_{Ft}L_{Ft}) \\ p_{Mt} &= Z^G_{Mt}\Psi_2(.)f^G_1(Z^G_{Mt}E_{Mt} + Z^G_{Ft}E_{Ft}) \\ p_{Ft} &= \frac{Z^G_{Ft}}{(1+\tau^O_t)}\Psi_2(.)f^G_2(Z^G_{Mt}E_{Mt} + Z^G_{Ft}E_{Ft}) \end{aligned} \quad (2.6)$$

With perfect substitutability, the relative wages between men and women in the *good jobs* and another occupations are given by:

$$\begin{aligned} \frac{w_{Ft}}{w_{Mt}} &= \frac{Z^G_{Ft}}{Z^G_{Mt}}\frac{1}{(1+\tau^G_t)} \\ \frac{p_{Ft}}{p_{Mt}} &= \frac{Z^O_{Ft}}{Z^O_{Mt}}\frac{1}{(1+\tau^O_t)} \end{aligned} \quad (2.7)$$

Labor Supply: All high-skilled workers supply labor inelastically to either *good jobs* or other occupations[7], with S_{gt} denoting the number of high-skilled individuals of each gender at date t for $g = \{M, F\}$. They differ in their ability a when performing tasks in *good job*, so that a worker with ability a earns $a * w_{gt}$. The distribution of ability a is

[7]In this model, we abstract the possibility of an individual be unemployed or choose not to work. In Cortes, Jaimovich and Siu (2018), they show a model extension considering the occupational choice and participation choice, with no substantial changing results. According to them, the model is robust to differences in the trend of women's participation.

gender and time-specific $a \sim \Gamma_{gt}(a)$, where Γ is the cumulative distribution function. By contrast, a worker who is employed in other occupations earns p_{gt} (since the ability to work in other occupations is similar to all high-skilled workers, it was normalized to 1).

The high-skilled worker has a choice to deal with. She will work in a *good job* if her ability is greater enough to make the earnings in this occupation $a * w_{gt}$ be bigger than the alternative in other occupation p_{gt}. So there is a cutoff level for ability a^*_{gt} that turns the individual indifferent between *good job* and other occupations, such that:

$$a^*_{gt} w_{gt} = p_{gt} \tag{2.8}$$

The worker whoose ability is $a < a^*_{gt}$ will choose to work in other occupations. Alternatively, if her ability is $a \geq a^*_{gt}$, he will choose to work in a *good job*. The fraction of workers of each gender who choose employment in the *good job*, Φ_{gt}, is:

$$\Phi_{gt} = 1 - \Gamma_{gt}(a^*_{gt}) \tag{2.9}$$

Equilibrium: The high-skilled labor market clears when the demand for labor input equals supply:

$$\begin{aligned}
L_{Ft} &= S_{Ft} \int_{a^*_{Ft}}^{\infty} a\Gamma'_{Ft}(a)da \\
L_{Mt} &= S_{Mt} \int_{a^*_{Mt}}^{\infty} a\Gamma'_{Mt}(a)da \\
E_{Ft} &= S_{F_t}\Gamma_{Ft}(a^*_{Ft}) \\
E_{Mt} &= S_{Mt}\Gamma_{Mt}(a^*_{Mt})
\end{aligned} \tag{2.10}$$

Given the number of high-skilled workers (S_{gt}), the effective labor in *good job* is the weighted ability conditional on being above the endogenous cutoff a^*_{gt}. In the other hand, the employment in the other occupations is the cumulative distribution function itself up to a^*_{gt}.

2.4.1 Analysis of model predictions

In this section, we analyze the model's predictions described above and rationalize the factors that drove the different gender trends observed in data between 1994 and 2017.

Using demand Equations (2.7) e (2.8), we have:

$$\frac{a^*_{Mt}}{a^*_{Ft}} \frac{Z^O_{Ft}}{Z^O_{Mt}}(1 + \tau^G_t) = \frac{Z^G_{Ft}}{Z^G_{Mt}}(1 + \tau^O_t) \tag{2.11}$$

Letting Δ be the percentage change between any time t and the imediately previous period $t-1$, Equation (2.11) can be written as:

$$\Delta a^*_{Mt} - \Delta a^*_{Ft} = \Delta \left(\frac{Z^G_{Ft}}{Z^G_{Mt}} \right) - \Delta \left(\frac{Z^O_{Ft}}{Z^O_{Mt}} \right) + \Delta(1+\tau^O_t) - \Delta(1+\tau^G_t) \tag{2.12}$$

Since $\Phi_{gt} = 1 - \Gamma_{gt}(a^*_{gt})$ is the likelihood of a high-skilled women (when $g = F$) or men (when $g = M$) is working in a *good job* at time t and a^*_{gt} is is the minimum ability to do a worker to sort into good job, then $\Delta a^*_{Mt} - \Delta a^*_{Ft}$ denote the diferential change in selection into the *good job* between men and women.

In Section 2.3, we saw that the results in Table 1 shows that ϕ_{F_t} has increased in the past years, which imply $\Delta a^*_{F_t} < 0$. At the same time, ϕ_{M_t} has fallen since 1994, which suggests $\Delta a^*_{M_t} > 0$. In other words, the data says that high-skilled women have progressively sorted into *good job* over time, whereas high-skilled men have increasingly sorted into other occupations. So the left-hand side of Equation (2.12) is positive.

For Cortes, Jaimovich and Siu (2018), equation (2.12) gives two channels for raising the women employment in good job. They can be seen on the right-hand side of the equation. The first one is related to the greater increase in the demand for female labor than male labor, which they call *Female Bias*. The growing demand for women $(\Delta(Z^G_{Ft}) > \Delta(Z^G_{Mt}))$ implies $\Delta(\frac{Z^G_{Ft}}{Z^G_{Mt}}) > \Delta(\frac{Z^O_{Ft}}{Z^O_{Mt}})$, assuming constant the demand for other jobs. In this case, there would be a female comparative advantage for this to happen. The second channel is the fall in discrimination parcel in *good jobs* relative to other occupations $\Delta(1+\tau^O_t) > \Delta(1+\tau^G_t)$. The second channel is quite difficult to analyze empirically, so we focus on the first one. To explore the *Female Bias*, we have to work with two hypotheses:

Assumption 1: The value of social skills has increased at the Brazilian labor market recently, such as Deming (2017), and Weinberger (2014) argued to have happened in the United States, and Borghans, Weel and Weinberg (2014) indicated to have also occurred in Britain ad Germany.

The principal difficulty in investigating it empirically in Brazil is the absence of an adolescent survey about young skills such as National Longitudinal Survey of Youth (NLSY79) in the United States to match with adult outcomes or job advertisements indexes as provided by Atalay et al. (2018). Without such information, the estimation of shifts in demand for social skills becomes hard to perform.

The hypothesis about the growing importance of social skills takes into account the difficulty of technology in substituting interaction between humans. The ability to lead, coordinate, and motivate people is more important nowadays than ever. Moreover, having empathy for others, interpreting what someone else might be thinking or feeling can help to reduce conflicts and raise an appropriate workplace to work. The spread of preset rules to perform many tasks have increased the efficiency of the labor market. Still, the coordination of the new kind of work is what improves efficiency. According to the model of Deming (2017), social skills reduce the transaction cost of combining tasks within a team in the output production and, consequently, raise productivity. Thus, the demand for occupations with a high level of social skills intensity has increased in the labor market.

Hypothesis one is reasonable to assume for Brazil when we think technology advance has similar impacts around the world. The process of replacing some jobs with machines takes little different time from a country to another, but in more extended periods, it tends to be alike. Internal rules could make some substitution harder to put into practice in some countries, like what we see with bus ticket collectors in Brazil. Some municipalities, such as Novo Hamburgo (Rio Grande do Sul state), for example, prohibit the substitution of bus ticket colletors for ticket validation machines by law. Besides, there exists some collective understanding that this replacement is unfair and has not to occur.

Even so, most job replacements by computers/machines have been widespread in the last two decades as bank tellers, online sales (began with airplane ticket sales and spread to e-commerce), the process of operating systems in enterprises, etc. The impacts of those changes tend to be similar to all countries.

Assumption 2: Women have a comparative advantage in social skills.

One potential source of female bias in *good jobs* is the women's comparative advantage in tasks requiring social and interpersonal skills. We follow psychology and neuroscience literature that confirms this hypothesis (see, for instance, Gilligan (1993), Hall (1978), Chapman et al. (2010), Baron-Coen, Knickmeyer and Belmonte (2005)) and take this as given for the Brazilian reality.

Assuming both assumptions, we relate the second channel of raising women employment in a *good job* (i.e., *Female Bias*) to the observed movements in the occupational skill requirements. First, we analyze if social skills are correlated to a higher female share in occupations, in the sense of natural female choosing for professions that demand higher social skills. Second, we investigate if there was an increase in demand for social skills since 2000 and whether it has attracted more women for occupations where the growing demand has occurred. Lastly, we examine whether social skills are, in fact, an comparative advantage for women, so they have helped raise women's wages.

2.5 DATA

In this paper, we use microdata from the RAIS database for the years 1994 and 2017, made available by Ministry of Economy - Special Secretary of Social Security and Labor[8].

[8]Special Secretary of Labor is the successor of Ministry of Labor and Employment (Ministério do Trabalho e do Emprego - MTE).

RAIS database compiles all formal employment information of Brazil, once all private and public companies have to send data to the Ministry about active employment contracts with information such as admissions, terminations, number of hours hired, salaries, and sociodemographic data such as age, education level, gender, among others.

The benefits of using the RAIS data come from the fact that it is not a sample of the workers in Brasil. Instead, it can be seen as an annual formal labor market census given the obligation imposed on Brazilian companies to report information about their employees on a regular basis. It covers almost all of the formal sector[9]. On the other hand, the disadvantage of it is the restriction to the formal sector, ignoring workers in the informal sector. So we have to point the results of this paper may underestimate the number of workers who perform manual and routine tasks once a considered part of them are in the informal sector.

Our analysis focus on full time (more than 36 worked hours), full-year (more than 12 months in the same job), 18-65 years old employees working in the private sector. We exclude individuals who work in armed forces or farming/forestry/fishing occupations. We do so for international comparability of results, once major studies exclude these occupations from the sample. We use the nominal wage of each year to obtain real wages deflating by consumer price index (INPC[10]). We construct real wages per hour using the number of hours hired for the work variable. We also use variables of gender, age, educational attainment, and time at work.

Even though we are working with real wages per hour, we exclude part-time workers to prevent these from being a source of wage bias. We also exclude workers from the public sector because of the public-sector wage premia, which can also be a source of bias. The downside of ignoring the public sector is that we may underestimate the women's probability of working in a good job because the female share is higher in the public sector. We consider the possible bias may have greater effects than the underestimation of the college-educated female population.

Our approach follows Autor, Levy and Murname (2003) and assumes the labor market clears when workers match to jobs that require the skills they have. For instance, a worker employed as an engineer has skills for math (cognitive) while a carpenter has skills for manual tasks. We work with CBO, which has two versions. The first one dated back to 1994 and remained until the 2000s. Another version of CBO was launched in 2002 and is still valid today. For our proposal, we made both versions compatible with each other with a procedure we describe in the next section.

[9]The RAIS information is not precise for the public sector.
[10]Índice Nacional de Preços ao Consumidor.

2.5.1 1994-2017 CBO Compatibility

Our first challenge was to make occupations from 1994 and 2017 compatible. To do so, we use the official crosswalk available on the Special Secretary of SocialSecurity and Labor's website. There were missing correspondents to the occupations with final code 90 in the 5-digit of the 1994 classification (from now CBO1994), which referred to the "other occupations", whose share in total employment was not negligible.

A solution we found was to use the crosswalk between CBO1994 and International Standard Classification of Occupation (ISCO) from 1988 (henceforth ISCO88), provided by Concla-IBGE[11]. We then constructed a text-mining tool, ran in R, to extract the CBO2002 correspondent of the ISCO1988 related to the specific missing CBO1994 from the PDF files. To solve the problem of existing a lot of CBO2002 correspondents to each ISCO1988, we use the Stringsim function from R to compare the text-similarity of the occupation's definition with the JW method. We combine this specific rule with our subjective analysis to determine the approximated correspondent.

There were additional occupations CBO1994 without a correspondent in CBO2002, beyond those we just mentioned (5-digit ended with 90). We deleted the CBO1994 with initial 2-codes equal to 99 because their registration identified "undefined occupations". Also, we excluded the CBO1994 not referenced in the CBO1994 book and the registrations in newer occupations with too few workers (occupations did not exist in recent years nor 1994).

Finally, we discard the information about occupations that no longer need to declare information to RAIS[12] as directors without employment relationship (for which FGTS[13] is not collected), legislators, domestic workers, interns, autonomous and cooperative workers. For comparison purposes, we also aggregate some CBO2002 that was count as a single occupation in 1994. As an example, we aggregate all of the occupations referred to as directors. The list of occupations we had to deal with this situation and the exclusions cited above is in Appendix A.1.

2.5.2 Measuring task content of Brazilian occupations (CBO)

Our second challenge was to discover a measure of the occupation's task content for CBO. The seminal work of Autor, Levy and Murname (2003) classifies the task content of occupation using the U.S. Department of Labor's Dictionary of Occupational Titles (DOT). The five measures of tasks (nonroutine analytic, nonroutine interactive, routine cognitive,

[11] Available in https://concla.ibge.gov.br/images/concla/documentacao/ibgexcbo94.xls

[12] The 2011 RAIS manual turned the declaration not mandatory to some occupations.

[13] FGTS is the Portuguese acronym for "Guarantee Fund for Length of Service", an obligatory social insurance fund, which is composed by the saving of 8% of the worker's earnings monthly to support them in case of specific eventualities,such as long-term sickness or resignation.

routine manual, and nonroutine manual) are defined aggregating five variables of DOT in the corresponding 3-digit of U.S. census occupation classifications (SOC). *Nonroutine analytic* corresponds to “GED Math (MATH)”; *Nonroutine interactive* corresponds to “Direction, Control, Planning (DCP)”; *Routine cognitive* corresponds to “Set Limits, Tolerances, or Standards (STS)”; *Routine manual* corresponds to “Finger Dexterity (FINGDEX)”; and *Nonroutine manual* corresponds to “Eye Hand Foot Coordination (EYEHAND)”.

Autor, Katz and Kearney (2006) used DOT as well, but collapsing the original five task measures into three task aggregations: abstract, routine and manual tasks. *Abstract* correspons to the simple average of two DOT variables: “MATH” and “DCP”. *Routine* corresponds to the simple average of “STS” and “FINGDEX”. Finally, *Manual* corresponds to the DOT variable “EYEHAND”. Autor and Dorn (2013) used the same measures.

Subsequent literature adopted the successor of DOT, the Occupational Information Network (O*NET). O*NET is a survey that questions a random sample of U.S. workers in each occupation about many issues concerning their abilities, skills, knowledge, work context, and work activities. The questions are rated on the ordinal scale. The O*NET survey began in 1998 and is updated frequently.

There is no such survey about the Brazilian labor market. So the way we find to capture the task content of CBO was to translate the O*NET scales to the international classification ISCO2008 and then translate to CBO. Our implicit assumptions are that the CBO’s task content is similar to SOC’s task content.

We map the O*NET-SOC occupational classification scheme to ISCO-08 coding following the step-by-step crosswalk provided by Hardy, Keister and Lewandowski (2018). Once we were able to collapse by mean O*NET scales in the ISCO-08 occupations, we use the same R text-mining tool as when we make the compatibility of CBO1994 - CBO2002 to extract the CBO2002 correspondent of the ISCO1988 related to O*NET-SOC.

O*NET task measures used in this paper (translated to CBO) are composite measures of O*NET Abilities, Skills, Knowledge, and Work Context level, using level and context scales. O*NET database counts with hundreds of different measures of skills, but it is not possible to use all of them simultaneously because the estimations would not be precise on account of multicollinearity (BACOLOD; BLUM, 2010).

We closely follow Deming (2017) and Cortes, Jaimovich and Siu (2018) to define *Social Skill Intensity*. It is given by the average of four O*NET measures from module “Social Skill”: i) Skill of Social Perceptiveness (2.B.1.a); ii) Skill of Coordination (2.B.1.b); iii) Skill of Persuasion (2.B.1.c); and iv) Skill of Negotiation (2.B.1.d). The first measure is related to the skill of paying attention to others and having some degree of developed emotional intelligence. It is obtained through the question: “being aware of other’s reactions and understanding why they react as they so". The coordination measure captures the ability to fast adapt to other's actions and is obtained by questioning whether the worker is used to "adjusting actions in relation to other's actions" when performing its job.

The persuasion measure is about the ability to make someone do or believe something by giving them good reasons to do so. The survey question about this item is: "persuading others to change their minds or behavior". Finally, the negotiation measure extracts the capability to reduce the differences between team workers and "to put them on the same page" by the questioning about: "bringing others together and trying to reconcile differences".

We combine the definitions from Deming (2017) and from Autor, Levy and Murname (2003) to have our measure of *Cognitive Task Intensity*, averaging six measures of O*NET: i) Ability of Mathematical Reasoning (1.A.1.c.1); ii) Skill of Mathematics (2.A.1.e); iii) Knowledge of Mathematics (2.C.4.a); iv) Skill of Management of Financial Resources (2.B.5.b); v) Skill of Management of Material Resources (2.B.5.c); and vi) Skill of Management of Personnel Resources (2.B.5.d).

The three first measures account for the mathematical development required for a job. At high levels of them, workers are required to know advanced calculus (as engineers, statisticians, physicists, etc.), while at low levels, they have to know only the basics in math, such as arithmetic. The math ability is a measure not only concerned with the acknowledgment of it but by the occupation requirement of mathematical reasoning and the capacity to solve math logic problems. By the knowledge module, the question is: "whether the occupation requires knowledge of mathematics.". For ability and skills modulus, the questions are "the extent to which an occupation requires mathematical reasoning" and "whether the occupation requires using mathematics to solve problemas".

The other three measures refer to the skills of managing in general, like managing money, material, and people. They are obtained by the following questions about the importance of "determining how the money will be spent to get the work done and accounting for these expenditures";"obtaining and seeing to the appropriate use of equipment, facilities, and materials needed to do specific work"; and "motivating, developing, and directing people as they work, identifying the best people for the job".

Routine Task Intensity is given by the average of two measures, as in Deming (2017): i) Work Context of Degree of Automation (4.C.3.b.2); and ii) Work context of Importance of Repeating Same Tasks (4.C.3.b.7). They refer to the measure of "how automated is the job" and "how important is repeating the same physical or mental activities over and over, without stopping of performing this job".

Finally, *Manual Task Intensity* is given by the average of two measures as in Cortes, Jaimovich and Siu (2018): Ability of Multilimb Coordination (1.A.2.b.2); and ii) Ability of Speed of Limb Movement (1.A.2.c.3). They capture the motor aptitudes of body, in the sense of coordination the limbs ("the ability to coordinate two or more limbs - for example, two arms, two legs, or one leg and one arm - while sitting, standing, or lying down. It does not involve performing the activities while the whole body is in motion."), and quick response with and movement them ("the ability to quickly move the arms and legs.").

We use level scales for most task intensity items, which goes from 0 to 7, but the work context variables have scaled from 0 to 5. For this reason, we standardized each O*NET measure to have mean zero and standard deviation one, similar to what most of the literature has chosen to do[14].

Our tasks indexes are defined by a combination of two or more O*NET measures, depending on the index. For *Cognitive Task Intensity*, for example, we use six O*NET measures, while for *Manual Task Intensity*, we use only two. So, once we obtain each of the indexes (Social, Cognitive, Routine, and Manual), we standardized each of them to have mean zero and standard deviation one. We do so to have comparable indexes. Each of them is evaluated now by its standard deviation to the mean.

Most of the literature uses task information registered at a specific point of time, as it was the same for each occupation over time. As in Cortes, Jaimovich and Siu (2018), we analyze the change in the task content of the occupations over time by using different versions of the O*NET database. The task content difference between two or more years is viewed as a change in the demand for a specific task measure over time.

We present the top and bottom ten occupations with the highest and lower level of of each index in 2000 and 2017 in the Appendix A.2. The occupational task indexes are available in Sulzbach (2020).

2.5.3 Descriptive Statistics

Using the matched dataset, which incorporates RAIS data by CBO2002 and the task content of each occupation obtained through O*NET, we show the descriptive statistics of the high-skilled labor market in Brazil. Table 4 shows the main characteristics of individuals. When we look to total college-degree workers, we see an increase in the percentage of women between 1994 and 2017. On average, the workers were older in 1994 (37.7 years old) than in 2017 (37.3 years old). The average age of women is lower in this market and has grown between the periods analyzed, unlike men. The nominal wage earned by workers is still lower for women, but the difference shrunk over time. In the following estimations, we use real wages. The number of formal workers by region is quite steady over time, with more than half of them located in the southeast.

[14]See Acemoglu and Autor (2011) for details.

Focusing on our variablesof interest, Table 4 indicates that cognitive tasks are more required in general than social skills. However, there is a difference between genders. For men, cognitive is more relevant in both years, having reduced relevance between 1994 and 2017. For women, on the other hand, both cognitive and social requirements have increased over the years, but social media is most required in both years. Occupations in high-skilled market little require routine tasks. When the sample refers only to female workers, the routine became less required overtime, contrary to that of men. Manual is the least required task for college-educated workers, with a negative index for all in both years.

Table 4 – Descriptive statistics - 1994-2017 – High-skilled workers (college-degree)

Variable (mean)		**Total**		**Women**		**Men**	
		1994	**2017**	**1994**	**2017**	**1994**	**2017**
Age (years)		37.7	37.3	35.7	36.6	39.0	38.1
Women (%)		0.39	0.52	-	-	-	-
Wage (R$)		1,629	6,383	1,050	4,928	1,994	7,962
Experience (months)		102.3	74.0	90.4	69.4	109.8	79.1
Worked hours (hours)		41.9	42.5	41.8	42.4	42.1	42.6
Region	South (%)	0.15	0.17	0.15	0.17	0.14	0.17
	Southeast (%)	0.54	0.54	0.53	0.52	0.54	0.55
	North (%)	0.02	0.03	0.02	0.03	0.02	0.03
	Northeast (%)	0.10	0.11	0.11	0.13	0.09	0.10
	Midwest (%)	0.05	0.07	0.05	0.07	0.05	0.06
Indexes	Social (sd)	0.60	0.59	0.56	0.62	0.63	0.54
	Cognitive (sd)	0.72	0.65	0.42	0.55	0.91	0.76
	Routine (sd)	0.11	0.15	0.23	0.16	0.04	0.13
	Manual (sd)	-0.88	-0.81	-0.92	-0.89	-0.86	-0.73
Number of workers (thousand)		682	2,897	264	1,508	418	1,389

Source: Elaborated by the author from RAIS data, 18-65 year old employees from private sector with at least college degree.

2.6 EMPIRICAL STRATEGY AND RESULTS

In this section, we present the empirical strategy to investigate the model predictions in Brazil and its results. We first investigate the correlation between social skills and the occupation's share of women. We then evaluate the hypothesis of social skills helping increase women's earnings through time, contributing to the female comparative advantage.

2.6.1 Social skills and occupational women share

To address the first analysis, i.e., whether social skills are correlated to a higher female share in occupations, we regress the level of female share of employment within each 4-digit level occupation, controlling for other tasks:

$$\phi_j = \alpha + \lambda_1\theta_{sj} + \lambda_2\theta_{cj} + \lambda_3\theta_{rj} + \lambda_4\theta_{mj} + \epsilon_j \tag{2.13}$$

where ϕ_j denotes the share of women in each occupation in each year, α is a constant, $\theta_k j$ with $k \in (s, c, r, m)$ represent occupational mean task content where s denotes social skills, c denotes cognitive, r denotes routine, and m denotes manual. ϵ_j is the error term of the equation. We perform a weighted least squares to maximize the efficiency of parameter estimation, once data does not guarantee that each point provides equally precise information about the deterministic part of the total process variation, that is the standard deviation of the error term is not constant. We estimate Equation 2.13 for both years, 1994 and 2017.

The results are at the first column of Table 5. Occupations that require more social skills have a larger share of women among the workers. In 1994, one deviation from the mean in social index increase the occupational women share in 8.4 percentage points, which is significant at 1% level. The results remain when we control for the other tasks, as shown by column 2 of Table 5. In this case, the coefficient is nearly 7.2 percentage points and still significant at 1% level. On the other hand, cognitive and manual occupational requirements reduce women's participation. The coefficient's signals are in accordance with the findings of Cortes, Jaimovich and Siu (2018) for the American labor market.

The same exercise was drawn with 2017 data. Social skill tasks content of occupations increase the female share by 10.8 percentage points, statistically significant. Controlling for other tasks contents of occupations, the social skill coefficient increase women share in 8.1 percentage points and is still significant at 1% level. In 2017, cognitive and manual tasks remained negatively correlated to female share of occupations, and routine skill persists being insignificant.

The results show there is some suggested evidence about the natural female choosing for professions that demand higher social skills hypothesis.

Table 5 – Importance of occupational tasks in the female share of high-skilled employment

	1994	1994	2017	2017
	(1)	(2)	(3)	(4)
Social	0.084	0.072	0.108	0.081
	(0.0156)***	(0.0213)***	(0.0152)***	(0.0195)***
Cognitive		-0.090		-0.098
		(0.0190)***		(0.0174)***
Routine		-0.018		-0.007
		(0.0136)		(0.0125)
Manual		-0.103		-0.143
		(0.0152)***		(0.0138)***
Constant	0.266	0.284	0.302	0.323
	(0.0123)***	(0.0115)***	(0.0120)***	(0.0104)***
Obs.	412	412	425	425
R^2	0.066	0.219	0.105	0.347

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***) p<0.01. The dependent variable is the female share of occupational employment. The regression is performed by weighted least squares.

2.6.2 Wage Evidence

The final investigation is about the hypothesis of social skills helping increasing women's earnings through time, contributing to the female comparative advantage. We perform a two-stage estimation. In the first one, we estimate gender-specific wage premia of each occupation in each year by quantile regression. Then, we estimate the impact of social skills on the wage premia.

In the first step, we estimate gender-specific wage premia for each 4-digit occupation by regressing log hourly real wages at 18-65 years old individual worker at the private market on age (five groups[15]), education (five categories[16]), tenure, and tenure squared by quantile regression. The regressions are run for each gender, year (1994 and 2017) and occupation . The coefficients of the constant are taken as the gender-specific occupational wage premia. The formal equation of the estimation is:

$$ln(w_i) = \alpha + X_i\beta + \varepsilon_i \tag{2.14}$$

where $ln(w_i)$ denotes log hourly real wages, α is the constant vector, which gives us the estimation of wage premia, X_i is the vector of all individual characteristics as age, education, tenure and tenure squared, β denotes the coefficient vector of X_i, and ε_i is the error term.

We are interested in analyzing the relationship between the wages using the conditional median function, so we perform a semiparametric estimation to capture the outliers of regression with more robustness. The variance-covariance matrix of the estimators (VCE) is estimated by bootstrap[17]. For proper convergence, we excluded occupations with less than ten workers registered.

Once we had the gender-specific wage premia of each occupation, we analyze the impact of social skill on it in the second step. We regress wage premia of each gender and each year on social skills controlling for other tasks.

$$\Omega_j = \rho + W_j\delta + \varepsilon_j \tag{2.15}$$

where Ω_j represents gender-specific wage premia of occupation j, ρ denotes the constant, W is a vector of explanatory variables, which includes social skills, cognitive, routine and manual tasks, occupational female share, and interaction of social and cognitive tasks, and ε_j is the error term of the equation. We perform weighted least squares to maximize the efficiency of parameter estimation, once data does not guarantee that each point provides

[15]Consider age1=under 25 years old; age2=between 25-35 years old; age3=between 35-45 years old; age4=between 45-55 years old and age5=above 55 years old.

[16]Consider educ1 = illiterate; educ2 = incomplete elementary school; educ3 = complete elementary school; educ4 = high school; educ5= college degree or more.

[17]Booststrap repetitions = 30

equally precise information about the deterministic part of the total process variation, that is the standard deviation of the error term is not constant.

Table 7 presents the main results of the regression for women. The first two columns consider only the relation between social intensity and women's wages. Occupations with social index one standard deviation above the mean increase the wage per hour in 0.376 the wage premia (in log) in 1994 and 0.148 in 2017 with 1% level of confidence.

When we control for other tasks measures and female share of each occupation (Model 2), the contribution of social skills to the women wage fall to 0.137 and 0.131 for years 1994 and 2017, respectively. In this case, R-squared is higher. The key result is the steadiness of social skill contribution to wages between 1994 and 2017. In the same period, cognitive skills have become relatively less important since 1994, with 1% level of confidence, from 0.225 to 0.157. Also, the inclusion of female share in the regression shows occupations with a strong female presence tend to pay fewer wages. The negative impact of female share, however, has become less relevant over the years, as it has gone from -0.844 to -0.409. Even in the high-skilled labor market, there is a tendency of women concentrating on less-paying occupations, but the effect of this concentration in earnings is narrowing.

The literature also documented that social skills raise the productivity of cognitive workers (MCCANN et al., 2010; WEINBERGER, 2014; DEMING, 2017). Controlling for the interaction between social and cognitive skills, the contribution of social skills remains similar to Model 2, with similar coefficients and stable trend. The cognitive index is still losing importance, and the interaction between social and cognitive helps increase women's wages in both years.

On the other hand, social skills seem to have lost their effects on wages over time for men. Its coefficient was 0.267 in 1994, considering Model 2, while in 2017, it was not statistically different from zero. The cognitive task has also become relatively less important since 1994, which coefficient went from 0.198 to 0.167, which is statistically different from one year to another. Even for men's earnings, the higher female share in occupation has negative effects. That is, men working in occupations female-dominated earn fewer wages than those working in occupations not so concentrated by women. In 1994, occupations female-dominated paid 1.28 fewer wages. The negative effect fell to 0.49 in 2017. When we control for the interaction between social and cognitive, the declining importance of social on men wage premia prevails, the drop in cognitive is even more significant, and the combination of social and cognitive gains relevance in determining men's wages.

Table 6 – Female occupational wage premia – High-skilled labor market

	Model 1		Model 2		Model 3	
	1994	2017	1994	2017	1994	2017
Social	0.376	0.148	0.137	0.131	0.137	0.139
	(0.067)***	(0.028)***	(0.081)*	(0.038)***	(0.082)*	(0.039)***
Cognitive			0.225	0.157	0.220	0.148
			(0.069)***	(0.034)***	(0.073)***	(0.035)***
Routine			-0.113	0.018	-0.107	0.028
			(0.047)**	(0.023)	(0.048)	(0.024)
Manual			-0.087	0.054	-0.092	0.055
			(0.060)	(0.029)*	(0.060)**	(0.029)
Female Share			-0.844	-0.409	-0.836	-0.396
			(0.177)***	(0.091)***	(0.177)	(0.092)*
Social*Cognitive					0.011	0.028
					(0.054)***	(0.026)***
Obs.	223	330	223	330	223	330
R^2	0.123	0.074	0.295	0.267	0.292	0.264

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***) p<0.01. The dependent variable is the gender-specific wage premia of occupation. The regression is performed by weighted least squares.

Table 7 – Male occupational wage premia – High-skilled labor market

	Model 1		Model 2		Model 3	
	1994	2017	1994	2017	1994	2017
Social	0.382	0.148	0.267	0.060	0.179	0.064
	(0.063)***	(0.027)***	(0.068)***	(0.039)	(0.080)**	(0.039)
Cognitive			0.198	0.168	0.282	0.157
			(0.068)***	(0.035)***	(0.073)***	(0.035)***
Routine			0.076	-0.015	0.081	0.004
			(0.047)	(0.022)	(0.050)	(0.024)
Manual			-0.211	-0.043	-0.252	-0.037
			(0.062)***	(0.027)	(0.063)	(0.023)
Female Share			-1.283	-0.491	-1.148	-0.429
			(0.198)***	(0.089)***	(0.206)***	(0.084)
Social*Cognitive					-0.060	0.034
					(0.051)***	(0.023)***
Obs.	320	396	320	396	320	396
R^2	0.101	0.068	0.471	0.248	0.363	0.264

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***) p<0.01. The dependent variable is the gender-specific wage premia of occupation. The regression is performed by weighted least squares.

Social does seem to be a comparative advantage for women in the high-skilled job market. Its relevance in determining wages remained the same between 1994 and 2017 for women, while it fell for men. Besides, social increases cognitive skills returns, so that individuals with the combination of both tasks have higher salaries. This is true not only for women, whose impact of the combination has doubled between the years analyzed but also for men, whose effect has gone from negative to positive.

2.7 CONCLUSIONS

The increasing demand for high-skilled workers in recent years is well known. In this paper, we provide evidence of gender-specific divergent trends in the likelihood of a high-skilled individual working in a *good job*. The rising women's probability contrast with the declining likelihood of college-educated male employed in one of these jobs. We show it is not fully explained by the increase of women's supply labor in the high-skilled market (college-educated) neither to demographic and spacial characteristics.

The paper's principal assumption follows the international documentation about growing importance and demand for social skills in the high-skilled labor market. Taking this evidence as valid for Brazil, we investigate the role of social skills in the rising demand for women in this market.

We contributed to the literature by building a measure of task content of Brazilian classification of occupations (CBO), including the measure of CBO's social skill index. We assume that occupational task content is similar across countries, and match the CBO's task intensity with RAIS database.

We showed a positive relation between social skills and the female share of occupations, evidencing a natural choice of women for professions intensive in this kind of skill. We also provided results consistent with neuroscience literature that women have a comparative advantage in performing tasks that require social skills. The relevance of such skills in determining wages is higher for women than men (in 2017). Social skills also raise cognitive returns in both cases, having increased its effect between 1994 e 2017.

We contributed to the literature with such evidence about the relationship between social skills and women's share in occupations and their wages. Promising research can derive from this first step, considering the Brazilian labor market.

3 JOB POLARIZATION AND TASK PRICES IN THE BRAZILIAN LABOR MARKET

Abstract

In this paper we propose to solve a puzzle concerning the polarization of wages in Brazil. Literature has documented a downward trend of wage change throughout Brazil's distribution of earnings, which contradicts the international evidence of U-shape pattern – the expansion of high- and low-skill occupation's wages at the expense of middle-skill ones, in response to increased/decreased demand for such jobs. The essay proposes an alternative way to analyze the polarization in wages by estimating task prices using panel data at the individual level and innovating using each occupation's continuous indexes. The results show a marked increase in the return of cognitive tasks between the two periods analyzed (2002-2003-2004 and 2012-2013-2014), while the advance of manual task price was milder in the same period, and the return of routine tasks did not change. The findings suggest polarization in the wage structure in Brazil, as observed in several other countries.

3.1 INTRODUCTION

Over the past two decades, a growing literature has focused on understanding how technological changes have affected the labor market. The so-called job polarization, in particular, has been well documented in several developed countries. The phenomenon that middle-class jobs – requiring a moderate level of skills, like autoworkers' jobs – seem to disappear relative to those at the bottom, requiring few skills, and those at the top, requiring greater skill levels, has been attributed to technological change, typically the rapid adoption of computers. Since the seminal study by Autor, Levy and Murname (2003) for the United States, many others have found evidence that, over the last years, both the level of employment and wages for jobs with a high degree of manual and cognitive tasks (those in the bottom and in the top of the wage distribution, respectively) have increased. Conversely, employment and wages in routine jobs (those in the middle of the distribution) have decreased in the same period[1].

The mechanism behind such a movement in employment and earnings is straightforward: the rapid technological progress expands job opportunities in both high-skill, high-wage occupations (cognitive) and low-skill, low wage occupations (manual) at the same time that reduces opportunities in middle-wage, middle-skill white-collar and blue-collar

[1] See, for example, Autor, Katz and Kearney (2006) for the United States, Spitz-Oener (2006) for Germany, and Goos and Manning (2007) for the United Kingdom. An extensive literature review is performed in Section 3.2.

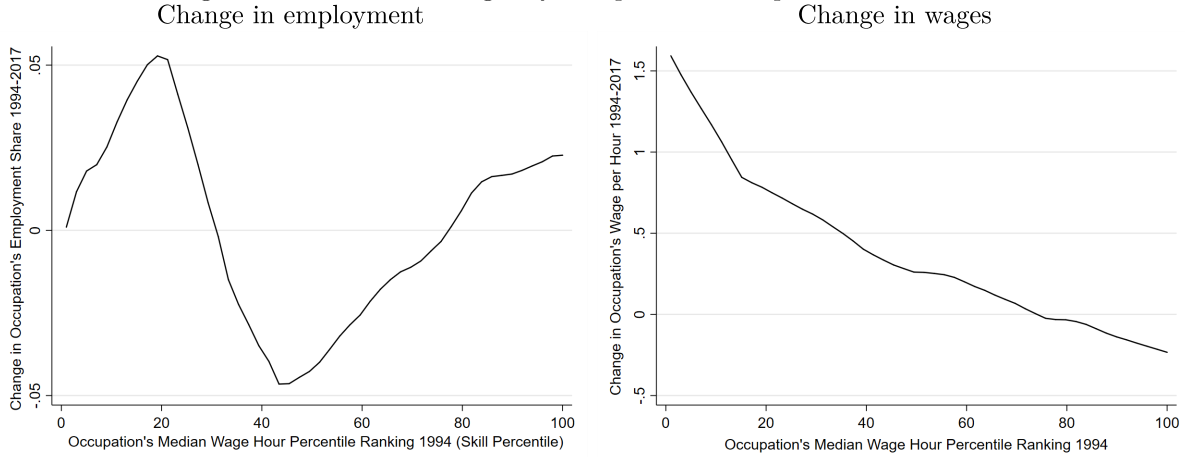
jobs (routine). Acemoglu and Autor (2011, p. 20) explain that, while the wide adoption of computer technologies decreases the demand for routine employees, it increases the demand for those whose productivity benefits from a more technologically advanced environment: "The rapid, secular price decline in the real cost of symbolic processing creates enormous economic incentives for employers to substitute information technology for expensive labor in performing workplace tasks. Simultaneously, it creates significant advantages for workers whose skills become increasingly productive as the price of computing falls".

Although the job polarization pattern is well established among developed countries, there is no consensus in Brazil. Some of the studies related to the literature report an increase in the demand for non-routine and cognitive tasks, and a consequent decrease in occupations with routine tasks over the last years, indicating job polarization (FUNCHAL; JUNIOR, 2013; RIVA, 2016; MALONEY; MOLINA, 2016; CORSEUIL; POOLE; ALMEIDA, 2018). However, others did not present evidence of the phenomenon in the Brazilian labor market (BULLA, 2014; MEDINA; POSSO, 2010; HERDEIRO; MENEZES-FILHO; KOMATSU, 2019). Nevertheless, except by the study of Riva (2016), there seems to be little divergence about the downward trend of wage change throughout Brazil's distribution of earnings, which contradicts the international evidence of U-shaped behavior.

By using data from the Annual Report on Social Information (RAIS[2]), we build two graphs that illustrate the literature findings on the Brazilian labor market. We follow the standard approach of studies in this field and compute smoothed changes estimated by locally weighted regression. Besides, we rank the occupations by their place in the 1994 wage per hour distribution along the horizontal axis. The occupation's wage refers to the median wage of all workers. Figure 2 shows an increase in employment (left graph) in low-skill occupations (at the bottom of the wage distribution) and high-skill occupations (occupations at the top of the wage distribution) between 1994 and 2017, which it is in line with the international evidence.

[2]Relação Anual de Informações Sociais - RAIS.

Figure 2 – Smoothed changes by occupational skill percentile 1994-2017



Source: Elaborated by the author from 1994-2017 RAIS data. Notes: The figure plots log changes in employment shares and occupational median wage per hour by 1994 occupational skill percentile rank using a locally weighted smoothing regression. Skill percentiles are measured as the employment-weighted percentile rank of an occupation's median log wage in RAIS dataset. Consistent occupation codes (CBO codes) for 1994 and 2017 were used using method decribed in subsection 2.5.1 of this thesis's first essay.

Contrary to what happens with employment, there is no evidence of polarization when it comes to wages of Brazilian workers (right graph of figure 2). The most substantial variations occur at the bottom of the wage distribution, whereas the smallest occurs at the top tail, showing a monotonic (decreasing) behavior. Over the 90s and 2000s – and, in particular between 1994 and 2017, the period analyzed in figure 2 – there was a huge increase in the use of computer-based technologies worldwide, including Brazil. Thus, the demand for workers with cognitive skills – those associated with high wages, in general – was expected to raise as it happened in other countries. Why are not the changes in wages similar to those in employment in the Brazilian labor market? What explains the singularity of the Brazilian case?

The significant rise in the minimum wage that occurred in the first decade of the 2000s may help to explain the lower tail of the curve in the right graph of figure 2, but it is useless when it comes to an understanding the behavior in the upper tail. Another possible explanation is that, despite of the large increase in the demand, there was an even higher increase in the supply of cognitive labor. This would explain why employment raised, whereas wages dropped in this period. To this happen, however, a large increase in human capital accumulation of Brazilian workers would have been necessary, a hypothesis that can easily be rejected.

A more promising alternative is based on understanding the changing balance among sectors. Literature usually focuses on the mean or median wage of occupation, but the average earnings may not accurately transmit the information about selection or composite prices on labor. Employment shifts might result in temporary imbalances on the worker's average characteristics of each occupation. A sector's employment growth may, for example, attract entrants with lower quality (skills), education, or experience, which reduces the average wage of the occupation. Consequently, the increase in the demand for a specific occupation may cause a fall in average wages paid for it. This suggests that an approach based on *price tasks* rather than the average or median wage of each occupation is more suitable for studying issues such as job polarization, as suggested by Cavaglia and Etheridge (2017) and Cortes (2016).

In this paper, we follow that latter approach and estimate price tasks in the Brazilian labor market. We use a panel data set that follows formal workers over time (from 2002 to 2014) thus enabling us to take account of their job changes. In addition to the novelty of using individual-level data, this paper employs the continuous task intensity index, presented in the first paper (chapter 1) of this thesis, which builds a proper correspondence between CBO and O*NET. The framework we build in this paper allows us to distinguish occupational average wages from prices on labor (selection or composite free prices). Our findings show that, when one uses task price, polarization emerges in terms of earnings. Similar to what is found by the literature for other countries, technological change increased earnings – measured by task price – in manual and cognitive tasks, and decreased in routine ones, over the period studied.

In terms of magnitude, we find that, in the first period (panel with years 2002-2003-2004), wages were 1% higher for individuals who work in an occupation that requires a unit of standard deviation of the mean of cognitive tasks, controlling for many variables, while in the second period (panel with years 2012-2013-2014), the effect of cognitive task on wages jumped to 3.2%. Moreover, we find a slight increase in the price of manual tasks, going from -0.3% to 0.2% between the two periods. On the other hand, routine tasks have not experienced valorization in the period, which suggests that Brazil's wage structure has been experienced a polarization. These findings are in line with what has been observed in several other countries – and are contrary to other papers on the Brazilian labor market, which found a decreasing trend.

Our approach is based on the hypothesis that there is poor correspondence between broad groups of occupations in Brazil, considering Brazilian Occupational Classification (CBO), and task intensity measures. We show that this is indeed the case and conclude that the classification suggested by Acemoglu and Autor (2011) and followed by most of the literature, in which occupations are classified in broad groups (1-digit occupation code), is not a good proxy for job tasks in Brazil. This suggests that the continuous index we adopt in this paper is preferable to CBO broad groups for analysis concerning job polarization. A further advantage of this index is that it takes into account the information of all tasks required in each occupation. For instance, it allows us to obtain the precise contribution of manual, routine, and cognitive skills for a manager rather than just classifying the occupation in one of these three categories. To the best of our knowledge, this paper is the first to use individual-level panel data to study task prices in Brazil and the first to find that there exists wage polarization in the Brazilian labor market.

Besides this introduction, this paper is organized into seven other sections. In the next one, we discuss the literature on task prices and job polarization, and how our paper fits into it. In the third section, we show that CBO classification does not provide a good approximation of job tasks in Brazil. In section four, evidence of the advantages of using continuous indexes is presented. In particular, we find suggestive evidence that by assigning values of abstract (cognitive), routine, and manual tasks for each 4-digit occupations of CBO, the polarization of earnings in the Brazilian labor market emerges. We then present the empirical strategy that we use to estimate task prices on labor in section five. Section six describes the data sample and the details of how task intensity is built. The results are discussed in section seven. Section eight concludes.

3.2 RELATED LITERATURE

The extensive literature on job polarization started with the seminal work of Autor, Levy and Murname (2003), who found the first evidence of polarization in employment in the United States and the first to emphasize the need for understanding the *task content* of occupations. Their main findings suggested that technological change causes an increase in employment and earnings of low and high ability workers, those who perform cognitive and manual tasks (occupations on the low and the high tails of the wage distribution), to the detriment of medium ability workers, who perform routine tasks (those in the middle of the wage distribution). Consequently, they show the relation between demand for skills and wage distribution was no longer monotonic as affirmed by the so-called *Skilled-Based Technological Change Model – SBTC*, that prevailed so far. The pattern was U-shaped instead.

The SBTC literature separated high-skill workers (with at least college education) and low-skill workers (others) and looked only at the impact of technological change on these two groups, assuming that there existed a one-to-one map between tasks and skills[3]. Latter empirical studies corroborated the findings of Autor, Levy and Murname (2003) of the U-shaped behavior such as Autor, Katz and Kearney (2006). For other countries, Spitz-Oener (2006), Goos and Manning (2007), and Goos, Manning and Salomons (2009) found the same pattern for Germany, the United Kingdom and for European Countries, respectively. Aditionally, Autor and Dorn (2013) showed that the change in the lower tail of the distributions of employment and wages in the U.S. labor market was mostly due to

[3]Some of the works of SBTC are, for the labor market in the United States, Katz and Murphy (1992), Autor, Katz and Krueger (1998), Goldin and Katz (2008) among others. For other countries, Murphy and Romer (1998) analyzes the United States and Canada, Kramarz and Lemieux (1999) include France in the analysis, for example.

a specific occupation category, namely services.

An alternative model to study job polarization was introduced by Acemoglu and Autor (2011). Their *Task-Based Model - TBTC* follows the ideas of the Ricardian model of exchanges and allows the attribution of skills for each task to be given endogenously. Its main implication is that technological changes do not translate into higher wages in all employment categories. Instead, there is an asymmetric effect on workers' earnings, such that some benefit to the detriment of others. According to this model, there is no fixed parity between skills and tasks, so that workers of varying skill levels can perform a given task, but those more qualified tend to have comparative advantages in performing more complex tasks.

Based on TBTC, other studies analyzed the phenomenon of polarization of labor markets, such as Goos, Manning and Salomons (2014) and Ikenaga and Kambayashi (2016), which performed analysis for sixteen European countries and Japan, respectively. Also, some studies emphasize the *offshorability* of tasks as a causal component of job polarization. The principal idea is that many medium-skilled employees were being hired from other countries, with effects on the domestic demand for these jobs and, consequently, on their returns (BLINDER, 2007; BLINDER; KRUEGER, 2008; GROSSMAN; ROSSI-HANSBERG, 2008; FIRPO; FORTIN; LEMIEUX, 2011; KELLER; UTAR, 2016; REIJNDERS; VRIES, 2017).

The scarcity of papers on job polarization and analyses of the demand for specific skills in the Brazilian labor market contrasts with the developed world. As we show in the next section, the complexity of finding which tasks (or skills) are required for each occupation – and in what intensity – is a plausible explanation for the difficulties the authors face. Nevertheless, a small number of studies have approached the subject by resorting to the available data. For example, Bulla (2014) applied the methodology presented by Goos, Manning and Salomons (2009) and used occupational wages as a proxy of the job's skill content for the period from 2002 to 2012. Data was obtained from the National Household Sample Survey (PNAD[4]) database. The author grouped occupations into six categories, following Acemoglu and Autor (2011), except for splitting out Routine cognitive and Non-routine cognitive in four other categories. Thus, the broad classification of occupations employed by that paper is based on the 1-digit code of CBO rather than on the content information extracted from occupations specific surveys as DOT or O*NET, as we do in this paper.

The main findings of Bulla (2014) is that employment declined for medium skill occupations, but grew slightly in low and high skill occupations, such that it is not possible to confirm the hypothesis of job polarization. Furthermore, the author found no confirmation of wage polarization as well. Earnings grew in the lower tail of the

[4]Pesquisa Nacional por amostra de Domicílios, carried out by Brazilian Institute of Geography and Statistics (IBGE[5])

distribution but fell in the upper tail, where the highest skill occupations are concentrated. Medina and Posso (2010) found similar results. The authors analyzed Brazil, Colombia and Mexico's job markets by following the *Task-Based Technological Change - TBTC* approach. Although the period studied is different (1981 to 2001), they also found no evidence for job polarization in Brazil.

Herdeiro, Menezes-Filho and Komatsu (2019) reproduced the SBTC model based on Tinbergen (1974) to identify the wage evolution by qualification groups of workers from 1981 to 2015. However, the authors did not investigate the effect on task, like other strands of the international literature have done following the *Task-Based Model*. They found that the shape of the demand curve for middle-skilled workers was equivalent to an inverted U shape, and showed that groups presented elasticity of substitution during the period studied. Also, Herdeiro, Menezes-Filho and Komatsu (2019) results showed a reduction trend in the relative wage between the groups. The *SBTC* is limited to explain earnings differences between just those two groups, skilled and unskilled, different from de the view of *TBTC*.

Other studies have found evidence of polarization when it comes to employment. Funchal and Junior (2013), for instance, found that the reducing barriers to computer entry in 1994 (end of the Computer Law) increased the demand for jobs that perform non-routine tasks, those which are complementary to computers and, on the other hand, reduced the use of occupations whose tasks are routine (which can be replaced). More recently, Corseuil, Poole and Almeida (2018) showed that digital technology adoption shifted the demand for skills toward increased use of non-routine and cognitive tasks. The authors used RAIS data from 1999 to 2006, period in which the provision of internet services expanded as a result of the privatization process put in place. Moreover, contrary to the previous literature, they used an occupational task content measure matching Brazilian Classification of Occupations (CBO) and the U.S. Department of Labor's Occupational Information Network (O*NET). The results showed that the labor market regulations ended up benefiting skilled workers, particularly those employed in non-routine and cognitive tasks (in contrast with the regulation policy intention).

Finally, Riva (2016) investigates the hypothesis of polarization in Brazil from a natural experiment, the end of the Brazilian market reserve policy on microcomputers (October 1992). The author found employment shifted "toward non-routine manual and away from routine tasks". Moreover, Maloney and Molina (2016) indicate potentially – and subtle – polarizing forces in Brazil and México after analyzing countries in Latin America, Asia, the Middle East, and Africa.

In this paper, we are less interested in proving the existence of employment polarization in Brazil, once there are solid literature and evidence (Figure 2) suggesting potential polarization forces. Our main interest is to investigate the peculiar pattern of wages in Brazil reported by Herdeiro, Menezes-Filho and Komatsu (2019) and Bulla (2014) and indicated by RAIS data (Figure 2). We match administrative information (RAIS data) and the task context of each occupation imported from O*NET questions, similar to what was performed by Corseuil, Poole and Almeida (2018). However, we adopt (following literature) different measures for cognitive, manual, and routine tasks and work with individual-level rather than city-industry levels. Furthermore, we innovate by using a continuous index of task content, which allows us to take all the information about the occupations' requirements.

3.3 THE NEED FOR A TASK INTENSITY MEASURE IN BRAZIL

The studies focused on individual-level in the context of labor market polarization used to aggregate occupation in task sectors by grouping 1-digit occupation as suggested by Acemoglu and Autor (2011). The authors showed that there is little difference in classifying occupations by the US Department of Labor's Dictionary of Occupational Titles (DOT), subsequently replaced by the Occupational Information Network (O*NET), and proxying job tasks by directly working with Census and Current Population Survey (CPS) 1-digit occupational categories.

Because of the precise correspondence between DOT/O*NET and data from CPS, Acemoglu and Autor (2011) were able to find that the intensity of the use of non-routine cognitive ("abstract") tasks from DOT/O*NET is highest in managerial, professional, and technical occupations, and lowest in service and laborer occupations. So the authors used Census and CPS occupational categories such as: (i) *Non-routine cognitive tasks* corresponds to managerial, professional and technical occupations; (ii) *Routine cognitive tasks* corresponds to sales, clerical and administrative support occupations; (iii) *Routine manual tasks* corresponds to production, craft, repair, and operative occupations; and (iv) *Non-routine manual tasks* corresponds to service occupations[6].

The classification mentioned above can be detailed in terms of the skills required in each one. Abstract (or cognitive) tasks are activities that require problem-solving, intuition, persuasion, and creativity, which are more often performed by occupations like medicine, science, law, engineering, management, among others. Routine cognitive tasks require some degree of analytics but also follow precise and well-understood rules. These tasks involve organizing, storing, retrieving, and manipulating information. Routine manual involves repetitive production and activities that can be specified series of instructions. Lastly, non-routine manual tasks are activities that require situational adaptability, visual and language recognition, and in-person interactions.

As we have seen, such aggregation has been widely adopted in the literature. Some

[6] Include Protective Service, Food/Cleaning Service and Personal Care

prominent studies which focused on individual-level are Cortes (2016), which use the Panel Study of Income Dynamics (PSID) data to analyze the United States labor market, and Cavaglia and Etheridge (2017), which use data from British Household Panel Survey (BHPS) and from German Socio-Economic Panel (GSOEP) to analyze the British and German labor markets, respectively. Both papers classified occupations in broad groups based on the categories used by Acemoglu and Autor (2011), i.e., through aggregating by occupational codes.

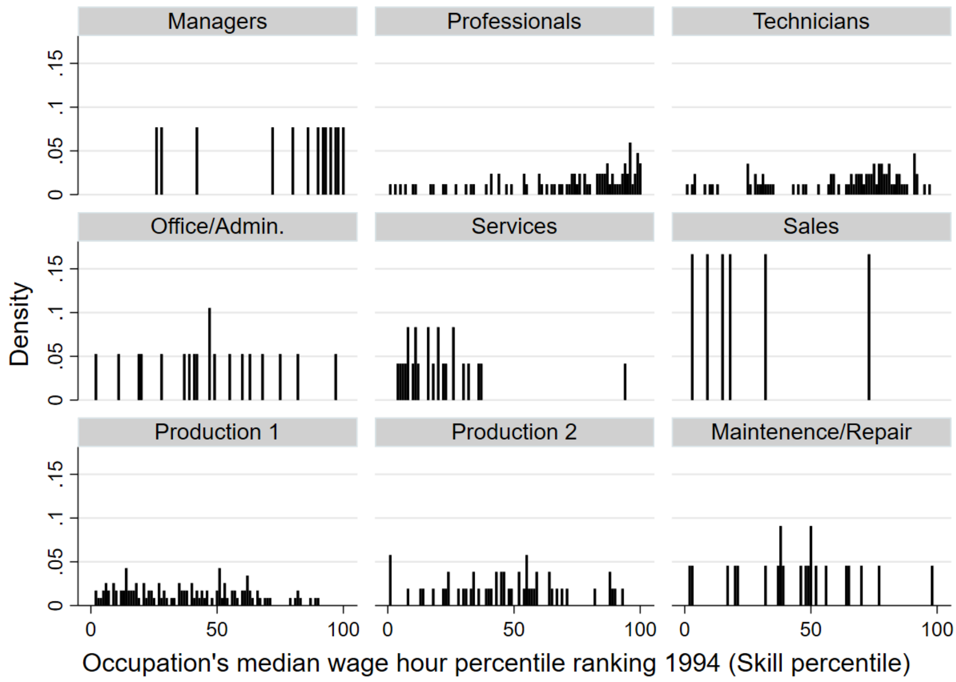
While the aggregation proposed by Acemoglu and Autor (2011) provides a good proxy for job tasks in developed countries, the same is not true for Brazil. The simple aggregation of occupations does not fit properly with the task intensity because the Brazilian Occupational Classification (CBO) focuses more on the specificity of the occupation functions than on the skills required by them (REIJNDERS; VRIES, 2017). The primary difficulty related to CBO category is differentiating production workers performing simple tasks from those in the same CBO who perform tasks that require higher skills, as supervisors, for example. It is not straightforward to classify some 1-digit CBO into non-routine cognitive, routine or non-routine manual groups.

For example, among the occupations considered routine by the standard classification of literature, such as "Workers of the production of industrial goods and services" (CBO 1-digit: 7 - Production 1 in Figure 3), there is a description of non-routine cognitive tasks (usually supervisors), workers with routine cognitive tasks (machine operators and electronics segments that in many cases need to have at least technical education) and, finally, routine manual work (e.g., assemblers, construction workers, and so one). In this case, we have workers from a broad occupation group in various positions in the distribution of wages.

Figure 3 shows how broad CBO categories (first digit[7]) are distributed among 1994 income percentile (log of the median hourly wage of each occupation considering four digits).

[7]The occupations are grouped by one digit. The only exception is CBO Group 5 divided into two 2-digit subgroups: 51 refers to services, and 52 to sales.

Figure 3 – Wage distribution by Acemoglu's groups of occupation – Brazil



Source: Elaborated by the author from 2002-2014 RAIS data, considering men between 25-65 years old working in the private sector. Notes: Occupational groups are aggregated by 1-digit of CBO, as in Acemoglu and Autor (2011). The only exception is CBO Group 5 divided into two 2-digit subgroups: 51 refers to services, and 52 to sales. The figure shows the probability density function of 1994 income percentile ranking (log of the median hourly wage of each occupation considering four digits) of each group.

Polarization literature assumes that the first three occupational groups (i.e., managers, professionals, and technicians) are at the top of the wage distribution, while office/administrative, sales, production 1, production 2, and maintenance and repair are in the middle of the distribution wages. Finally, the last group (services) is in the lower tail of the wage distribution.

As we can verify in Figure 3, this division does not accurately reflect reality in Brazil. Although services are concentrated in the lower tail of wage distribution, as pointed by literature, other broad CBO groups do not exactly follow the logic of the aggregation. For example, the group of workers in the broad category "sales" seems to be more concentrated in the lower of wage distribution rather than in the middle, as polarization models predict. Also, the "production 1" group appears to have a concentration in both lower and middle tails, in disagreement with the model's predictions (which advocate this group has concentration only in the middle of the distribution of wages). Even groups as "professionals" and "technicians" are not segregated in the upper tail.

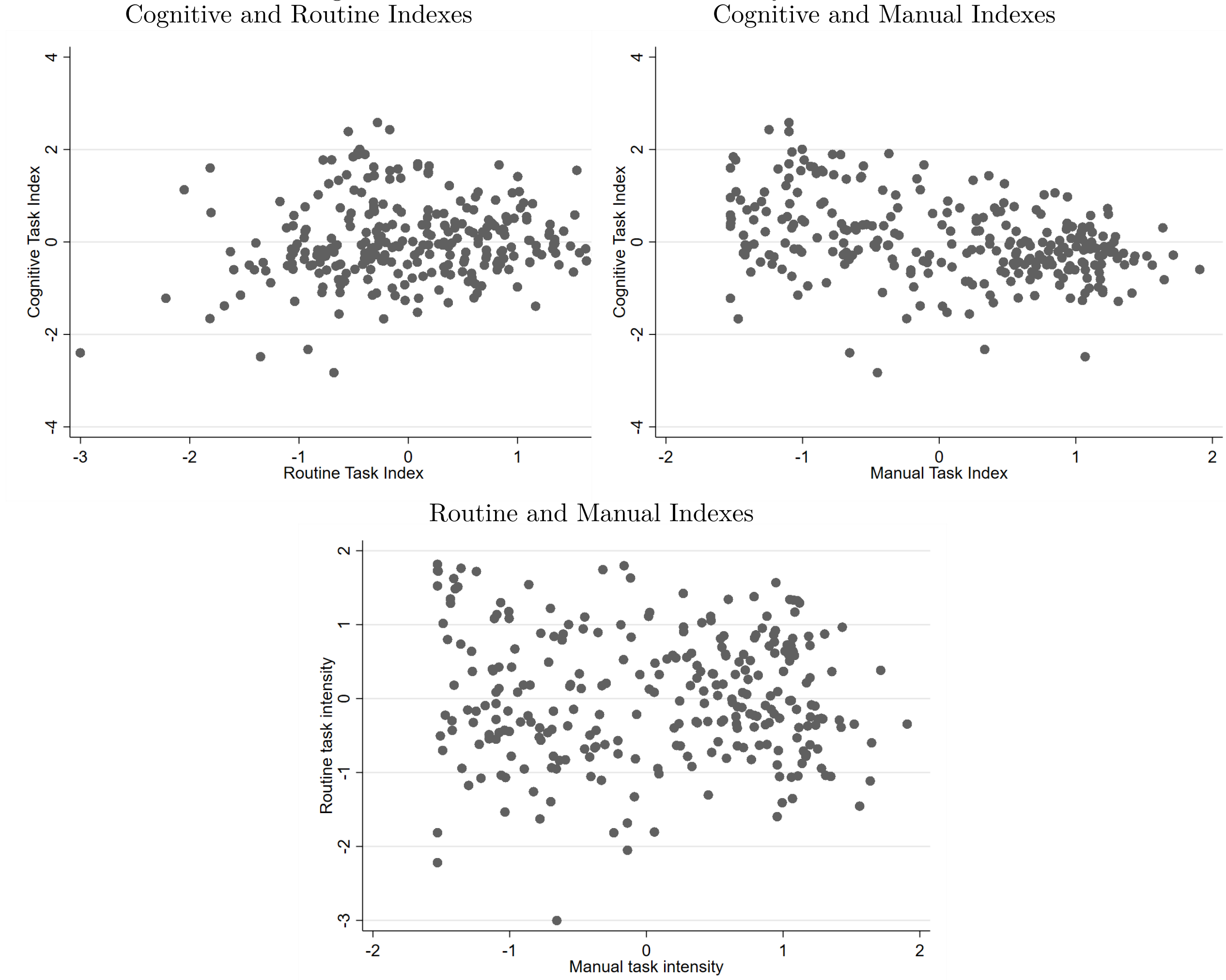
The second reason why the simple aggregation of occupations does not fit properly with the task intensity of occupations in Brazil is the loss of relevant information on task pricing when we divide occupations into exclusionary groups. The tasks required by each occupation are not exclusively cognitive, or mainly routine or manual. For example, the medical profession requires a high level of cognition and, at the same time, some degree of manual skill. There might be a high level of correlation between the tasks required by each occupation.

Figure 4 shows that there is no apparent negative relationship between the tasks required by occupations. The highest levels of cognitive intensity tasks required by occupations are not necessarily related to the lowest level of manual or routine task requirements, and so on. If we separate occupations into segregations groups, therefore, we lose information about task requirements.

As our interest in this paper is to evaluate each task's return, it is essential to take into account every piece of information about occupational task requirements. To understand the importance of a method that takes into account task intensity, consider the following example. An occupation requiring both cognitive and manual tasks, such as physicians, has experienced increased returns. In this case, such an increase can be due to either task or to both. In either case, the intensity of manual and cognitive tasks must be considered. Clearly, when one takes into account only the cognitive aspect of the job, there might be a substantial loss of information on the job market dynamics.

We overcome the difficulty mentioned above by using the continuous indexes for task classification of occupations built-in paper 1 of this thesis. The next section presents suggestive evidence that polarization can emerge once we employ such indexes.

Figure 4 – Relation between task intensity indexes



Source: Elaborated by the author from 2002-2014 RAIS data, considering men between 25-65 years old working in the private sector. Notes: The figure presents the correlation between occupational task intensity indexes.

3.4 EVIDENCE BY USING CONTINUOUS INDEXES

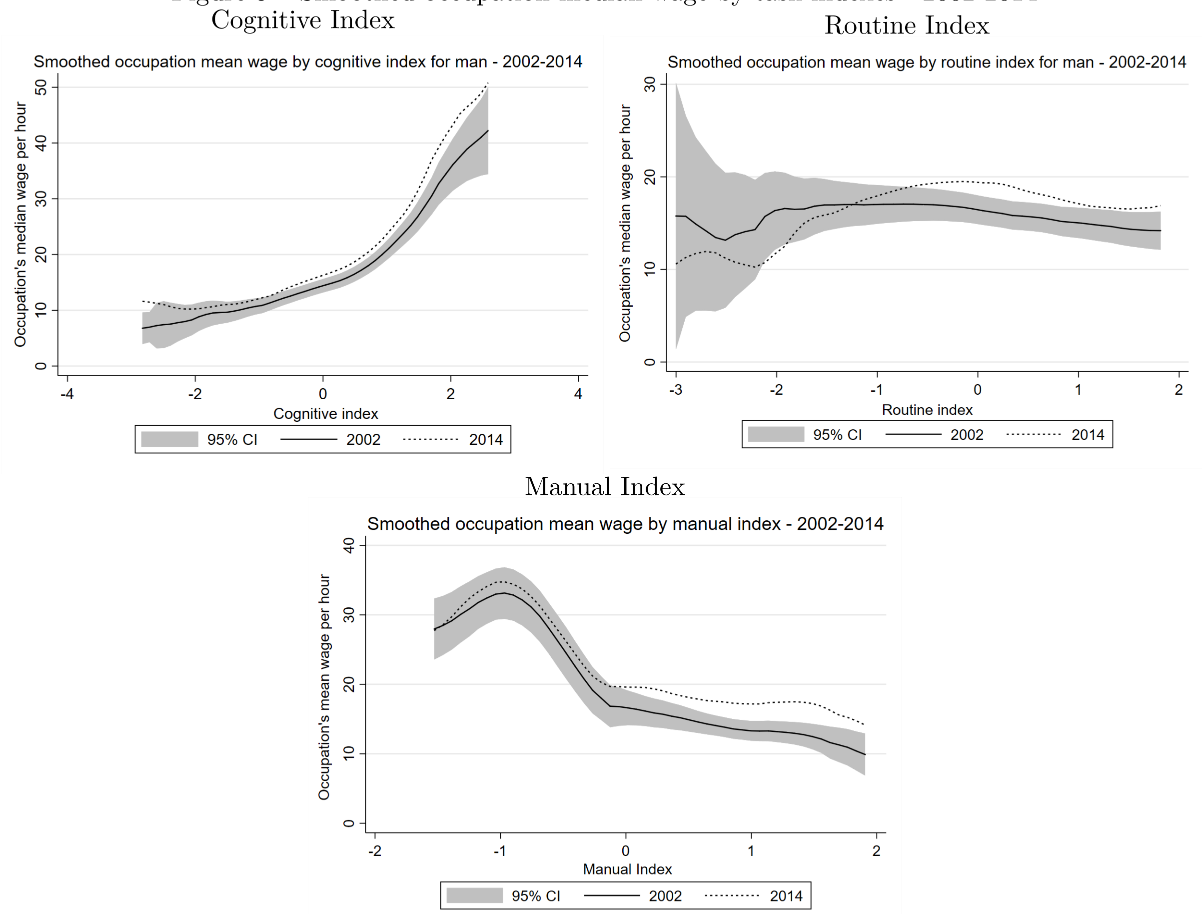
By using continuous indexes, evidence of polarization in the wage distribution shows up, contrasting to Figure 2 (right side). We use data from 2002 to 2014 once it is the period contemplated by panel data we analyze in the next chapter. Similar results are found for the period 1994-2017, which we explore in the appendix.

Figure 5 shows an increasing relationship between the cognitive index and occupation's median wages, as expected. The comparison between 2002 (solid line) and 2014 (dashed line) suggests an increase in income between periods for occupations of all levels of cognitive request, with more marked growth in occupations requiring higher cognitive ability levels. Still, it was within the confidence interval.

The same analysis for manual task requirement suggests that occupations that require less manual skill are better paid than those which require higher levels of this task (the negative relationship between return and manual index). In the comparison between 2002 and 2014, however, the highest hourly wage growth occurred in occupations that require higher levels of manual (from positive deviations from the average), according to what literature has recently shown.

Concerning the routine task, we do not see the same monotonic behavior found for the other two cases (positive for cognitive and negative for manual). Routine task returns are increasing up to levels near zero (no deviation from the mean) and decreasing after that. The comparison between 2002 and 2014 suggests that the most significant increase in wages was in occupations requiring routine levels between 0 and 1 standard deviations from the average.

Figure 5 – Smoothed occupation median wage by task indexes - 2002-2014



Source: Elaborated by the author from 2002-2014 RAIS data, considering men between 25-65 years old working in the private sector. Notes: The figure presents the occupational median wage per hour of each task index, and each year using locally weighted smoothing regression.

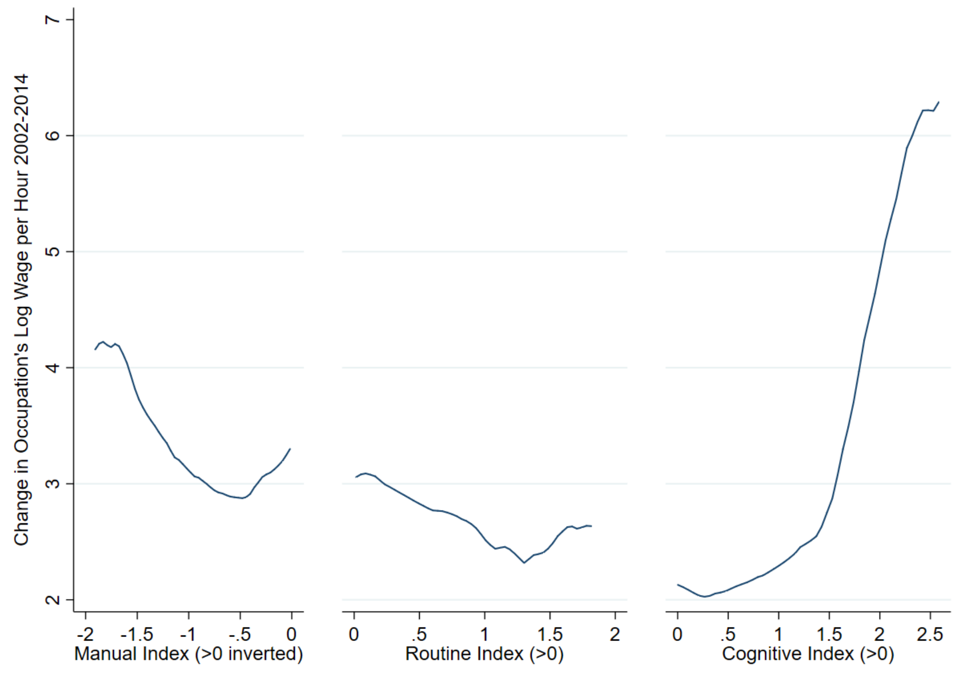
From the Figure 5 above, we see the return of occupations requiring a high degree of cognitive skills increased in the analyzed period, as well as occupations requiring a high level of manual skills. This result is in line with what the literature found to major advanced countries. The main idea of the findings is technological changes led to an increase in the demand for workers with problem-solving tasks to complement the task performed by machines. At the same time, it also led to an increase in the demand for workers performing manual tasks which machines still cannot perform. The limitation of computers/machines of performing manual tasks is related not only to the lack of specific algorithms but also to the cost-benefit ratio of putting a machine to do what a human workforce does at low cost (in the case the algorithms are accessible). For example, the task of distributing goods on supermarket shelves and stores could be automated, but the gains from investing in it do not outweigh the costs since the return on this type of work performed by human beings is still low.

In such a context, despite the increase in demand for routine between 2002 and 2014, it did not happen in the same intensity as for other tasks. Routine was the most affected task by the incorporation of new technologies in the economy. For this reason, the return of higher levels of routine requirements is not only lower than the intermediate level, as it increases less in the period analyzed. The increase in the upper side of the routine index is next to the confidence interval, indicating there may have been little growth in the return to that level of the index.

If occupations could be perfectly segregated into exclusive groups, we would say the high levels of cognitive and manual tasks – both categories with significant wage improvements – lie at opposite ends of the occupational skill distribution. So, polarization would be a natural conclusion. However, as we see in Figure 4, there is no evidence that we are dealing with a linear measuring ruler. Although it is not common, there is some occupation requiring a high level of cognitive ability while requiring some average degree of manual task.

One exercise we are allowed to do is simulate a “linear rule” and see what pattern the change in wages would take. Figure 6 shows a possible U-pattern of wages in the Brazilian labor market, with a greater increase in cognitive tasks. The occupational log median wage per hour change between 2002 and 2014 lie on the vertical axis. Along the horizontal axis, we consider only positive task indexes (greater than zero). In the case of manual, for the sake of visualization, we inverted the index, so that -2 refers to occupations that require two standard deviations more manual tasks then mean and zero refers to the mean requirement of manual content. From Figure 6, one can observe that, whenever one considers the continuous task intensity rather than just the occupation’s wage percentile ranking as a proxy for skills, the downward trend found by the literature no longer holds. Instead, a kind of U-pattern appears, with more pronounced growth for the return of cognitive tasks. For more accurate conclusions, we have to estimate the underlying prices of tasks using panel data. The empirical strategy of doing that is described in the next section.

Figure 6 – Change in occupation's log median wage per hour by positive task indexes - 2002-2014



Source: Elaborated by the author from 2002-2014 RAIS data, considering men between 25-65 years old working in the private sector. Notes: The figure shows the relation between change in occupational log median wage per hour between 2002 and 2014 and the positive requirements of each task (task indexes greater than zero). In the case of manual, for the sake of visualization, we inverted the index, so that -2 refers to occupations that require two standard deviations more manual tasks then mean and zero refers to the mean requirement of manual content.

3.5 EMPIRICAL STRATEGY

In this section, we present the empirical strategy to estimate the underlying prices of tasks using panel data, which follows formal workers over time and enables us to take into account their job changes.

3.5.1 Econometric Framework

Our framework is based on Mincer's model focusing on the return of task prices and controlling for several variables. Wages depend on individual and time, such as:

$$\ln w_{it} = \alpha + \beta X_{it} + \lambda_1 \theta_{ci} + \lambda_2 \theta_{ri} - \lambda_3 \theta_{r^2 i} + \lambda_4 \theta_{mi} + \psi_t + \phi_i + \epsilon_{it} \tag{3.1}$$

where w_{it} denotes logarithm of real wage adjusted for hours worked of individual i in time t; λ is a vector of four coefficients of interest that measure price of each task; θ_c denotes cognitive task index, while θ_r represents routine task index, and θ_{r^2} means routine task index squared to capture nonlinear function of the task (Figure 9), and finally θ_m denotes manual index. The task indexes are common to each occupation (CBO 4-digit), so individual within the same occupation have the same index for each task. In the model specified by Equation 3.1, we assume task intensity indexes don't change over time. We use the 2014 O*NET information to construct the indexes.

In Equation 3.1, X_{it} is a set of observable characteristics capturing individual and time-specific factors; ψ_t denotes year dummies, controlling for any eventual trends in wages typical to all individuals; ϕ_i denotes the (unobserved) ability of individual i, and ϵ_{it} is an idiosyncratic error. We separate $\lambda\theta$ from βX in this expression precisely because of its importance. The vector of covariates (X_{it}) includes age (in level and a quartic polynomial), nine educational attainment dummies, the experience of worker i in time t (in months), nine dummy variables for all company sizes, dummy variables for activity sectors [8], and a dummy for whether the individual lives in a metropolitan area.

The empirical problem we have to deal with is the possibility of selection bias once the workers' choice of occupation with different task intensity is endogenous. In other words, the sectoral choice depends on the unobserved ability of worker ϕ_i. In general, econometricians need to find an instrumental variable to solve this kind of problem or assume a strong hypothesis of the functional form (CAVAGLIA; ETHERIDGE, 2017).

Using panel data, we can control for selection bias by estimating a fixed effect on the individual. For Botelho and Ponczek (2011), which analyze the wage differences between formal and informal markets in Brazil, even when there is selection bias, the fixed effect estimator is consistent since the unobservable variable ability (which can be interpreted as a set of unobserved characteristics as preferences, ability, and others) is constant in time. To achieve identification in this model, all coefficients of interest are estimated by running the regression using fixed effects at the individual level.

Identification (and, therefore, consistency) in this approach depends on the standard assumptions in fixed-effect regressions. We assume the unobserved variable ϕ_i is fixed over time, whats requires that unobserved human capital factors are fixed, so that changes to human capital are captured by the observable variables (CAVAGLIA; ETHERIDGE, 2017). Second, we assume weak homogeneity, that is, covariate variables (X_{it}) are not correlated to the idiosyncratic error ϵ_{it}. Formally, $cov(X_{it}, \epsilon_{it}) = 0$. Finally, we assume there are not perfect colinearity among variables.

Also, we draw attention to the fact that identification is only possible because there are individuals changing occupations and, therefore, changing the task intensity of the job performed over time. The results we present in this paper measure the differential of wages for those who move between occupations with different task intensities.

[8] We consider section aggregation of activities sectors of CNAE. For details of classification, see Appendix.

3.6 DATA

This article uses the RAIS database for the years 2012 to 2014, and compare to results from the database from 2002 to 2004. The Ministry of Economy - Special Secretary of Labor[9] provides RAIS, which compiles all formal employment information of Brazil, including private and public sectors. The information contains the number of os workers, numbers of hours worked, wages, and sociodemographic data such as age, education level, gender, etc.

For the panel data, we use a particular version of the RAIS database, provided by the Ministry of Economic - Special Secretary of Labor, to identify the same individual over time. Although RAIS's database has been publicly available online since 1985, this particular database is only available from 2002. We consider workers solely with active-link with a company (public or private) on December 31st of each year. In the case one worker has two or more formal jobs, we only consider the one with the highest wage reported. RAIS database reports a picture of the formal employment of December of each year. Changes in employment throughout the year are not contemplated. Once we are dealing with the formal sector, we discard information when the wage reported was lower than the minimum wage.

We use the nominal wage of each year and divide it by the number of hours hired to get the wage per hour variable. For the comparison between years (and panels), we focus on real wages, considering INPC price index (Índice Nacional de Preços ao Consumidor) available from Instituto Brasileiro de Geografia e Estatística (IBGE). Even though we are working with wage per hour, we exclude part-time workers to prevent these from being a source of wage bias. So, we focus on full time (more than 36 hours per week) and full-year (with active-link above 12 months) workers, and we exclude individuals who work in armed forces or farming/forestry/fishing occupations. We do so for international comparability of results, once major studies exclude these occupations from the sample.

We consider men in their prime age, between 25-65. The choice to focus on the male is due to the fewer econometric challenges when dealing with selection in the labor market. We also discard information about the public sector, and we do so for two reasons. The first is related to the difference in wages paid by each sector. Because there is a significative public sector wage premium, reported by World Bank (2019), we aim to avoid problems with selection. The second is related to the job stability of the public sector. Once we aim to discover the patterns of shifts in the labor market, the public sector's stability could generate biased estimators.

Finally, we exclude information about specific occupations which, according to the

[9] Special Secretary of Labor is the successor of the Ministry of Labor and Employment (Ministério do Trabalho e do Emprego - MTE).

RAIS Manual 2011 (MTE, 2012), should no longer provide information referring to each of its employees to the Ministry of Economy - Special Secretary of Labor through RAIS. It refers to occupations as directors with no employment relationship for whom it is not collected FGTS, autonomous workers, eventual workers, elective officeholders (governors, deputies, mayors, councilors, etc.), interns, domestic servants, and cooperative workers.

As discussed, we estimate Equation 3.1 using fixed effects at the individual level. We use two separate data panels, one for the period 2002-2003-2004 and the other for 2012-2013-2014. The fixed effect estimator controls for the time-invariant component ϕ_i by de-meaning the wage within each individual-spell. Also, it controls for the time-variant component common to all individuals ψ_t, which is captured by year dummies in each panel. The task price vector λ is estimated by the influence of each task intensity index on the individual wages. These prices are identified only as changes concerning a base year (2004 in the first panel and 2014 in the second one). In the results reported in Section 3.7, therefore, these coefficients should be interpreted as the changes in the return of task price over time, with respect to that in the year 2004 in the first panel and 2014 in the second.

We construct the occupational task indexes (θ) following the methodology of paper 1 of this thesis. The CBO2002 task indexes are composite measures of one or more O*NET Abilities, Skills, Knowledge, and Work Context variables. For Cognitive Task Intensity Index, we use six O*NET measures, while for Manual and Routine Task Intensity Indexes, we use only two. Most of the O*NET variables are defined based on questions rated on the ordinal scale, mostly from 1 to 7. However, some of them are rated from 0 to 5. So, for this reason, combined with the fact of each task are composite to different number o O*NET variables, we standardized each of task index (cognitive, routine and manual) to have mean zero and standard deviation one. We do so to have comparable indexes. Each of them is evaluated by its standard deviation to the mean.

3.7 RESULTS

In this section, we present the results of the estimation of the model previously discussed. Table 8 presents four model variations from Equation 3.1 for the both panels. The first column refers to a shorter version of the model, capturing only the impact of tree tasks on wages.

The second column differs from the first one by including a square version of routine tasks to capture the nonlinear effect of these specific tasks on wages seen in Figure 9. The third column refers to a broad model that includes all of the observable variables vector X_{it} and time fixed effect ψ_t, but not considering the nonlinear part of the routine task index. Finally, the fourth column adds to the previous estimation, the component nonlinear of the routine task index.

The results are for the subsample of males for excluding additional factors related to women's entry into the labor market. For robustness exercise, we estimate the same models with a full sample, including women.

Table 8 – Estimated Task Price

	2002-2003-2004 PANEL			
	I	**II**	**III**	**IV**
Cognitive	0.020	0.020	0.010	0.0010
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0001)***
Routine	-0.003	-0.002	-0.004	-0.003
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0001)***
Rout_sq		-0.003		-0.002
		(0.0001)***		(0.0001)***
Manual	-0.002	-0.003	-0.002	-0.003
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0001)***
Obs.	17,991,567	17,991,567	17,991,567	17,991,567
F	4,425.75	3,367.90	2,368.08	2,046.01
Prob>F	0.0000	0.0000	0.0000	0.0000
	2012-2013-2014 PANEL			
	I	**II**	**III**	**IV**
Cognitive	0.053	0.051	0.032	0.031
	(0.0002)***	(0.0002)**	(0.0002)***	(0.0002)***
Routine	-0.011	-0.005	-0.003	0.000
	(0.0001)***	(0.0001)***	(0.0002)***	(0.0002)*
Rout_sq		-0.009		-0.006
		(0.0001)***		(0.0001)***
Manual	0.002	-0.002	0.004	0.002
	(0.0002)***	(0.0002)***	(0.0002)***	(0.0002)***
Obs.	30,938,643	30,938,643	30,938,643	30,938,643
F	11,283.90	8,794.96	6,158.71	5,321.72
Prob>F	0.0000	0.0000	0.0000	0.000
	CONTROLS			
Considered	No	No	Yes	Yes

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***) p<0.01. The regression is performed controlling for selection bias by estimating a fixed effect on the individual in each panel and also controlling for observable variables as Age, Age quartic, Educational attainament (5 dummies), Experience, Company size (9 dummies), Sectors of activity (17 dummies) and Metropolitan area.

From Table 8, we see cognitive tasks have a higher return in all models. By model I, an individual who works in an occupation that requires a unit of the standard deviation from the mean of cognitive tasks have wage increased in 2.0%. Both routine and manual tasks have negative returns in this model (-0.3% and -0.2%). When we add routine index squared, the results are similar (2.0%, -0.2% and -0.3%, respectively), and its signal (negative) indicates the concave function for the wage concerning routine index according to what we had noticed by Figure 9. So, higher routine index increase wages, but at some point (and Figure 9 suggests it is in the central point, where the occupation requires routine index as the mean of all occupations) the wages start to decrease. All results are verified at 1% of significance.

Controlling by observable variables and year dummies (model III), we see the coefficients of cognitive and routine decrease to 1% and -0.4%, respectively, while the manual's coefficient remained steady. When we add routine index squared, the results are similar again (1.0%, -0.3%, and -0.3%, respectively), and the squared coefficient is -0.2%.

These findings are quite interesting, but they do not answer our main question: to check if there were advances in the return of specific tasks over time. To answer that, we have to compare these results to the same exercise done with panel 2012-2013-2014.

By the comparison between two periods (panels) of Figure 8, we see the most valorization happened to cognitive tasks, whose return went from 1% to 3.2% in the simple model considering controls (model III). In this comparison, routine returns increased marginally from -0.4% to -0.3%, but remains negative in the period, and manual returns increased from -0.2% to 0.4%. There seems to be a trend convergence with what happens in other countries in the world, with the valorization of cognitive and manual tasks and little difference in the routine returns. In Brazil, however, the more than proportional increase in the return of cognitive tasks draws attention.

When we consider the fourth column of Table 8 (model IV), the valorization of cognitive tasks remained similar to what we saw in the model III, going from 1% to 3.1%. In this model, routine returns increased from negative (-0.3%) to null (0.0%), but after the optimum point, its contribution remains negative and higher (-0.2% to -0.6%, respectively in panel 2002-2003-2004 and panel 2012-2013-2014). That is, an excess of routine task reduce wages. Manual tasks that contributed negatively to wages in the previous period (-0.3%), started to contribute positively (0.2%).

That is, it is good to have cognitive and manual (to a lesser extent) skills, and it is bad to have a lot of routine skills. Those who went from a lower level of routine to medium had a salary increase, but the return of occupations requiring a high degree of routine became even more negative.

3.8 CONCLUSION

Across most developed countries, technological change has made occupational structure shift towards a valorization of cognitive and manual jobs, with increases in both employment and earnings. On the other hand, routine intensive occupations had their employment levels and wages reduced as computers replaced workers. The final equilibrium movement has been called the polarization process because of the position of each task in the wage distribution — cognitive in the upper tail, manual in the bottom, and routine in the middle.

However, in Brazil, literature has found a pattern similar to the international evidence of polarization when we look at employment levels, but a different one when it comes to wages. In fact, the evidence by using CBO data suggest that there is a downward trend in wage change over time as we move to high-skilled jobs. A hypothesis that may explain this trend is that employment shifts might result in a temporary imbalance on the worker's average characteristics of each occupation, such that using occupational mean (or median) wages may not be the best way to assess the polarization phenomenon on the wage distribution.

In this paper, we estimate task prices in the Brazillian labor market by using panel data at the individual level. Once we can identify the individual over a period, we can extract the exact price of each task, controlling for any other individual's characteristics. We impute to individuals the task measures associated with their occupation. Thus, an innovation of this paper is to take into account the continuous indexes of each occupation instead of segregating them in broad exclusive groups as the massive literature has done. By doing so, we preserve information on occupational task requirements, especially in occupations with a high level of two of the three tasks. With all the information available, the task price estimate becomes more accurate.

We find a noticeable increase in the price of cognitive tasks between the two periods analyzed (2002-2003-2004 and 2012-2013-2014). An individual who works in an occupation that requires a unit of the standard deviation from the mean of cognitive tasks had wage increased in 1.0% in the first period jumping to 3.2% in the second. There was an increase in the price of manual tasks as well, going from -0.3% to 0.2%. On the other hand, routine tasks have not experienced valorization in the period, suggesting some polarization in the wage structure in Brazil, as observed in several other countries.

4 INFORMALITY, TASK CONTENT AND POLARIZATION IN THE BRAZILIAN LABOR MARKET

Abstract

In this paper, we shed light on the divergence trend in employment changes through the earnings distribution between formal and informal sectors by proposing a discussion about the impact of occupational task content on the probability of an individual being informal and the wage gap between the two sectors. The results show that, even after controlling for selection bias and observables characteristics, the probability of being informal is negatively correlated to cognitive and routine while the only positive effect among the three tasks comes from manual ones. Also, cognitive is the most important task explaining the gap both in 2003 and 2015. Its importance grew between the period, contributing to the maintenance of the wage gap over time. On the other hand, manual and routine tasks have little effect on it, so that we can conclude that tasks show a low power to influence the closing formal-informal wage gap.

4.1 INTRODUCTION

The phenomenon of job polarization has been well documented worldwide (see chapter 2 of this thesis). In particular, its main idea that technological improvements make the demands for cognitive and manual tasks (those which require high and low skills, respectively) increase whereas causes an opposite effect (decreasing demand) on routine tasks (middle-skills) has been shown to be true for developed countries and also for Brazil[1]. The baseline model of this literature, namely the *Task-Based Model*, distinguishes tasks from skill but does not take into account the potential effects of differences between formal and informal sectors. In fact, both this theoretical basis and the empirical literature which have been followed it do not devote much attention to the consequences of the migration from informality to formality.

A possible reason for the absence of studies on informality is that the vast literature on polarization has focused on advanced economies (developed countries), where the informal sector does not represent a significant share of the labor market. However, informality plays a substantial role in emerging economies, which suggests that this sector's presence may affect the results in terms of polarization of employment and earnings. In this paper, we address this issue in the context of the Brazilian labor market. We are

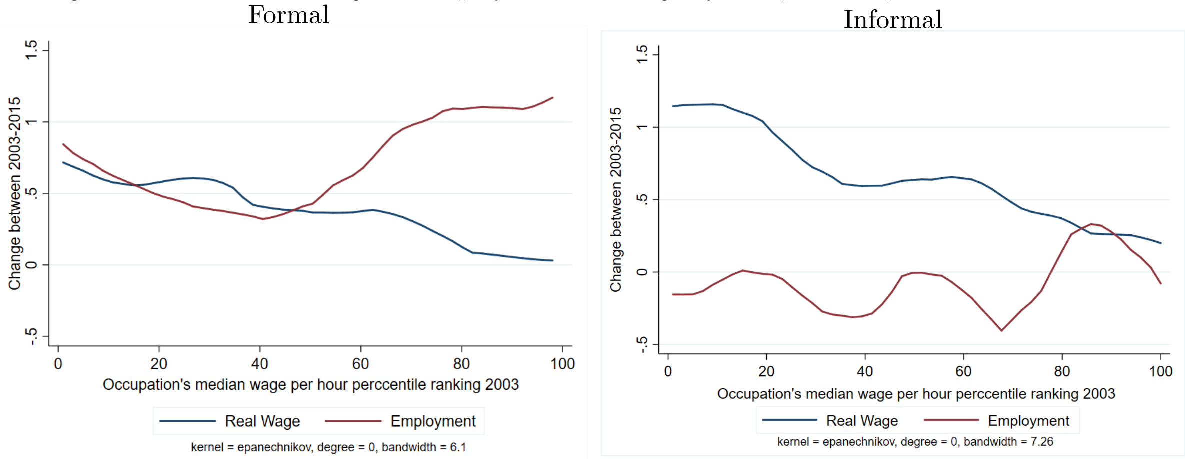
[1] In chapter 2 of this thesis, we present a survey of the literature, which brings the findings for countries such as United States, Japan, and United Kingdon. For Brazil, chapter 2 itself provides evidence that there is polarization in terms of earnings whereas the literature has already found polarization in terms of employment.

particularly interested in investigating the relationship of occupational task requirements and the extensive migration from informality to formality occurred over the 2000s.

Brazil is indeed a suitable case when one wants to investigate it. First, the informal sector represents a substantial share of the total of employment in the Brazilian labor market. Second, its importance has been decreasing over the last two decades. Due to the unprecedented process of employment formalization in the 2000s, non-registered employment (informal), which represented 40.1% of the total labor market in 2003, dropped to 28.2% in 2015, according to the National Household Sample Survey (PNAD). Such a variation provides an opportunity to explore the correlation of informality with occupational tasks in the context of job polarization.

Our motivation for including both sectors formal and informal into our analysis is reinforced by Figure 7. One can observe that the employment growth in Brazilian labor market, calculated between 2003 and 2015, presents different shapes according to the sector. While formal jobs seem to behave according to what the international literature founded (U-shaped), the same is not true for the informal ones. It is hard to define a pattern for this sector, but we can see that there was no increase at all, except for occupations between the 80th and 100th percentiles, which presented a low growth. An aggregate or joint analysis that includes both sectors is necessary to understand the whole dynamics of the labor market in Brazil. When one includes only the formal one – as the literature does –, there is a significant loss of information.

Figure 7 – Smoothed changes in employment and wage by occupational percentile - 2003-2015



Source: Elaborated by the author from the 2003-2015 PNAD dataset. Notes: The figure plots log changes in employment shares and occupational median wage per hour by 2003 occupational skill percentile rank using a locally weighted smoothing regression. Skill percentiles are measured as the employment-weighted percentile rank of an occupation's median log wage in the PNAD database. Occupattions are aggregated by tree-digits for sample matter.

In contrast, wage behavior looks the same in both sectors. The change in occupational wages between 2003 and 2015 is decreasing in both formal and informal sectors as we move to higher percentiles of the wage distribution, which contradicts the findings of the literature on polarization. However, as we saw in essay 2 of this thesis, when we look only to the average or median wage of each occupation – as the literature does –, we do not consider the task effect on wage growth. The employment growth in a specific occupation might induce a temporary reduction in wages due to the characteristics of new entrants, frequently less experienced and with lower education. These findings suggest the rising importance of analyzing the direct effects of tasks on wages. The Brazilian wage polarization "puzzle" was addressed in chapter 2, such that in this chapter we focus on the effects on the informality on employment.

By employing the correspondence between CBO and O*NET presented in the first chapter of this thesis, we are able to analyze the role of cognitive, routine, and manual tasks on employment and wage dynamics in both Brazil's formal and informal sectors. We investigate how the process of formalization in the Brazilian labor market has interacted with the changes in demand for different tasks – over the 2000's. Such an analysis allows us to answer the following questions. Does the task required by an occupation affect the probability of an individual being informal? Do tasks explain, in some way, the wage gap between formal and informal sectors? To the best of our knowledge, we are the first to include informality in the polarization framework and evaluate tasks shifting between the two sectors.

We estimate the contribution of each one of the three tasks on the likelihood of being informal by performing a three-stage probability model based on Heckman's selection model. The estimates from a structural probit showed that informality is positively associated with manual tasks and negatively with routine and cognitive ones even after controlling for selection bias and observables characteristics. In the period analyzed (2003-2015), while the routine's negative effect has halved and cognitive's became statistically not different from zero, the manual's effect has remained unchanged.

Our findings suggest the informality in the Brazilian labor market is highly associated with jobs that require higher levels of manual tasks and little levels of routine ones. This result provides elements to conjecture that the large decrease in the informality rate that the Brazilian labor market has witnessed over the last two decades might be explained by the migration of workers who perform manual tasks. In fact, given the nature of the informal market, such as the absence of labor contracts, poor access to credit markets faced by employers, and government persecution against firms, informal jobs are generally in services (e.g. sales, transportation, housekeeping). On the other hand, routine jobs are often related to mechanization, which is unlikely to be found in an informal sector because of the aforementioned characteristics. As we saw in chapter 2, there has been an increasing demand for manual skills in the formal labor market – a decreasing for routine ones –, which jointly with the process of labor formalization may explain the large rise in the lower bottom o the wage distribution. While salespeople, drivers, and housekeepers are likely to get a formal job, machine operators, manufacturing, among others, may be migrating to the informal sector or moving to a job requiring manual in the formal sector, as stated by literature.

We also found that the Brazilian labor market's formalization process was accompanied by a reduction in the wage gap between the two sectors. In 2003, the formal sector's average wage was 31% higher than the informal's. This difference decreased to 17% in 2015. We perform an Oaxaca-Blinder decomposition to verify the occupation task's effects on the wage gap and whether their relative impact has changed over time. Our findings show that cognitive tasks have an important role in the formal-informal wage gap and contribute to keeping the difference in favor of the formal sector over time. Manual and routine tasks have little effect on the wage gap, such that both show a low power to close such a gap. The main component in helping to narrow it is education, whose relevance is positive but decreased between 2003 and 2015. The formal sector valorization of cognitive tasks analyzed in chapter 2 of this thesis has worked for keeping the magnitude of the gap – or, at least, to make the decline in this difference slower over time.

The remainder of the paper is structured as follows. Section 2 presents the related literature. Section 3 analyzes the data used in this paper and presents evidence that the informality rate is decreasing in jobs that require cognitive and routine tasks and increasing in the levels of occupation's requirements of manual tasks. Section 4 sets out our econometric methodology for modeling the propensity of being formal and for performing Oaxaca-Blinder decomposition of the formal-informal wage gap. Section 5 presents our results. Section 6 concludes.

4.2 RELATED LITERATURE

Despite the existence of a vast literature addressing several informality issues, especially in developing countries where a large share of the labor market is informal, there is no consensus about its concept. Traditional views treat informality as a subsistence sector that acts as a buffer between formal employment and unemployment so that it rises during restructuring or recession (PORTES; SCHAUFFLER, 1993). One of the views suggested by Magnac (1991) follows the same idea: low skill individuals are kept out of the formal sector due to minimum wage laws, firing regulations and social security payments, among other labor market regulations. This view is in line with the idea of *survival strategy* for low-skill individuals and entrepreneurs who are too unproductive to achieve the formal sector (ULYSSEA, 2018).

The second view suggested by Magnac (1991) is based on a traditional microeconomic model of individual decision. It takes the comparative advantage idea from the model of Roy (1973), in which there is a competitive market where individuals rationally choose between formal and informal sectors. Their decision criteria take into account both the wage offered in each sector and their skill advantage. As jobs in different sectors may require different skills, wages reflect amenity differences. The two views explored by Magnac (1991), however, are focused exclusively on the supply side of the labor market (workers). The evidence from developing countries has shown that it is important to study the demand side as well.

As a response to the need for studies on the demand side, literature has broadened the informality concept by taking into account the firm's behavior. Similar to the "worker approach", there are two possible ways to interpret an informal firm. One the one hand, informality could be a "reservoir of potentially productive entrepreneurs who are kept out of formality by regulatory costs, most notably entry regulations" (ULYSSEA, 2018). On the other hand, the informal sector could be composed of "parasite firms entrepreneurs", which are competitive and take advantage of a no-cost payment to be more profitable. Either case, as explores by Ulyssea (2018), firms can exploit two margins of informality: (i) not register their business, the extensive margin, and (ii) hire workers "off the books", the intensive margin.

Once we have learned about the alternatives views on informality, we must investigate which of one is corroborated by the empirical literature. The first thing to notice is that the segmentation theory has been losing strength over time as papers have shown little adherence to empirical data (CARNEIRO; HENLEY, 2001; MENEZES-FILHO; MENDES; ALMEIDA, 2003). Some studies have documented transitions between the formal and informal sectors in both directions in Brazil, reducing the appeal for segmentation theory (CURI; MENEZES-FILHO, 2006; NERI, 2002).

Ponczek (2007) suggested there is no complete segmentation in terms of labor skill, once "there are firms in both sectors using the same type of labor". A similar result is reported by Meghir, Narita and Robin (2015), which presented evidence that low skill workers are seen in both sectors. Moreover, for a given productivity level, firms can choose to be formal or informal as well. For the authors, segmentation is endogenously determined by "interplay frictions, the institutional requirements for formal firms, and the penalties of informality". The above discussion suggests that the answer may lie between the two theories: a part of the informal sector workers face a segmented market whereas others can choose between working in the formal or informal market (FREIJE, 2001; TANNURI-PIANTO; PIANTO, 2002).

In this paper, we follow the assumption of endogenous segmentation of Meghir, Narita and Robin (2015) so that individuals rationally choose between sector based on the wages offered and their characteristics. We contribute to the discussion of informality estimating the role of tasks required by occupations in this decision. Our findings are consistent with those which have presented evidence of the presence of low-skilled workers in formal and informal sectors. We show that manual skills, for example, are associated with jobs in both sectors.

The literature on the subject has also faced the empirical difficulty of defining what characterizes an informal worker and an informal firm (ULYSSEA, 2006). Because the focus of this paper is on the supply side of the labor market, the discussion below is centered on workers. In Brazil, the existence of legislation that makes it mandatory for all workers to have a signed employment record book may help to overcome such a challenge. However, even in this case, it may not be clear whether one must consider a self-employee as an informal worker, for example. Several studies define the informal sector as the sum of workers who do not have a formal contract (those without an employment record book[2]) as well as those who work by themselves (self-employee). However, another strand of the literature defines it as the group of workers who do not contribute to the social security (ULYSSEA, 2006). In this paper, we follow the restrictive approach used by Menezes-Filho, Mendes and Almeida (2003) and consider as informal workers those who do not contribute to social security (which includes those who do not have a formal contract).

4.3 DATA

This section presents the data sample and the descriptive statistics of the Brazilian formal and informal labor market.

4.3.1 Data Sample

This study uses the National Household Sample Survey (PNAD) database (micro-data) for 2003 and 2015, carried out by Brazilian Institute of Geography and Statistics (IBGE[3]). Members of each household are asked questions concerning their labor activities and sociodemographic characteristics.

We consider individuals in their prime age between 25-65 and focus on full time (more than 36 hours per week) and full-year (with active-link above 12 months) workers. We drop the information about the public sector to reduce econometric challenges in dealing with selection bias and differences in wage composition between sectors. We discard information about individuals who work in armed forces or farming/forestry/fishing occupations and about individuals who produce for their consumption or work in construction for their use. We do so for international comparability of results, once major studies exclude these occupations from the sample.

We define informal employment as the absence of a signed employment record book[4]

[2] Trabalhadores sem carteira de trabalho assinada
[3] Instituto Brasileiro de Geografia e Estatística
[4] Trabalhadores sem carteira de trabalho assinada

or the case of self-employed or employers without contribution to the pension system[5]. We use the nominal income from the main job for persons over ten years in age, and we work with the logarithm of wage per hour, deflating by consumer price index (INPC[6]), available from Instituto Brasileiro de Geografia e Estatística (IBGE).

Our parameters of interest in this article are cognitive, routine, and manual task intensity of each occupation. To construct the intensity index to the occupations, we use the method of chapter 1 of this thesis, which map data from Occupational Information Network (O*NET) to the Brazilian Classification of Occupations (CBO[7]). We append the task intensity indexes with PNAD data.

4.3.2 Descriptive Statistics of Formal and Informal Labor Market

In this subsection, we present some descriptive statistics from the two sectors in Table 9 and show evidence of the effect of tasks on the informality rate in Figure 8. The sample characteristics (Table 9) indicate that cognitive tasks are more required in the formal sector in both years (2003 and 2015), despite having negative averages in both sectors. Once the majority of occupations in both sectors in the Brazilian economy do not yet require cognitive tasks, the negative signal was expected. The same is true for routine tasks: it is more demanded in the formal sector, but it has a positive average in the formal sector and negative average in the informal one. In contrast, manual tasks seem to be more demanded in the informal sector. In fact, its average in the informality increased from 0.18 to 0.28 between 2003 and 2015, while in the formal sector, the requirements of these tasks are negative over the period.

The percentage of women increased in both sectors between the two years analyzed and their presence remains greater in the formal market. The proportion of individuals nonwhite also increased between 2003 and 2015 but is bigger in the informal sector. The average age of individuals working in the formal sector increased from 34.3 years to 36.9 years between 2003 and 2015, while the informal sector's average age is slightly higher, changing from 35.6 years to 38.3 years. Education has also increased over time and remains higher in the formal sector. The average wage received by individuals in the formal sector was higher in 2003 than the average wage in informal and remained greater in 2015, but the difference has shrunk in the period analyzed.

[5] Conta própria ou empreagadores que não contribuem para o INSS.
[6] Índice Nacional de Preços ao Consumidor
[7] Classificação Brasileira de Ocupações - CBO

Table 9 – Descriptive statistics of formal and informal markets

	Formal Workers				**Informal Workers**			
	2003		2015		2003		2015	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Cognitive	-0.14	0.97	-0.11	0.97	-0.35	0.85	-0.38	0.80
Routine	0.08	0.94	0.01	0.95	-0.44	0.86	-0.51	0.80
Manual	-0.08	1.02	-0.10	1.03	0.18	0.92	0.28	0.88
Woman	0.34	0.47	0.38	0.48	0.27	0.44	0.28	0.45
Nonwhite	0.41	0.49	0.52	0.50	0.54	0.50	0.64	0.48
Age	34.30	10.64	36.85	11.36	35.58	11.79	38.34	12.39
Urban	0.97	0.18	0.96	0.20	0.95	0.22	0.92	0.27
Education	9.90	3.80	11.11	3.50	7.85	3.97	9.07	3.99
Tenure	67.78	77.95	73.66	86.35	70.51	89.64	86.06	102.60
Head of Household	0.55	0.50	0.49	0.50	0.59	0.49	0.53	0.50
More than 1 job	0.03	0.16	0.02	0.13	0.03	0.16	0.02	0.13
Other Income	0.06	0.24	0.03	0.17	0.06	0.23	0.02	0.15
Wage	854	1,277	1,971	2,678	540	779	1,434	2,162
Hours worked	45.34	5.69	43.74	4.67	46.44	6.57	44.15	5.50
Employee	0.87	0.34	0.85	0.36	0.48	0.50	0.44	0.50
Self-employed	0.07	0.25	0.10	0.30	0.47	0.50	0.52	0.50
Employer	0.06	0.24	0.05	0.22	0.06	0.23	0.04	0.20
Age begin work	14.92	4.10	15.79	3.79	13.89	4.25	14.75	4.02
Hourly Wage	8.94	14.15	10.83	14.59	5.30	8.47	7.68	11.05
Obs.	44,268		60,349		29,589		23,725	

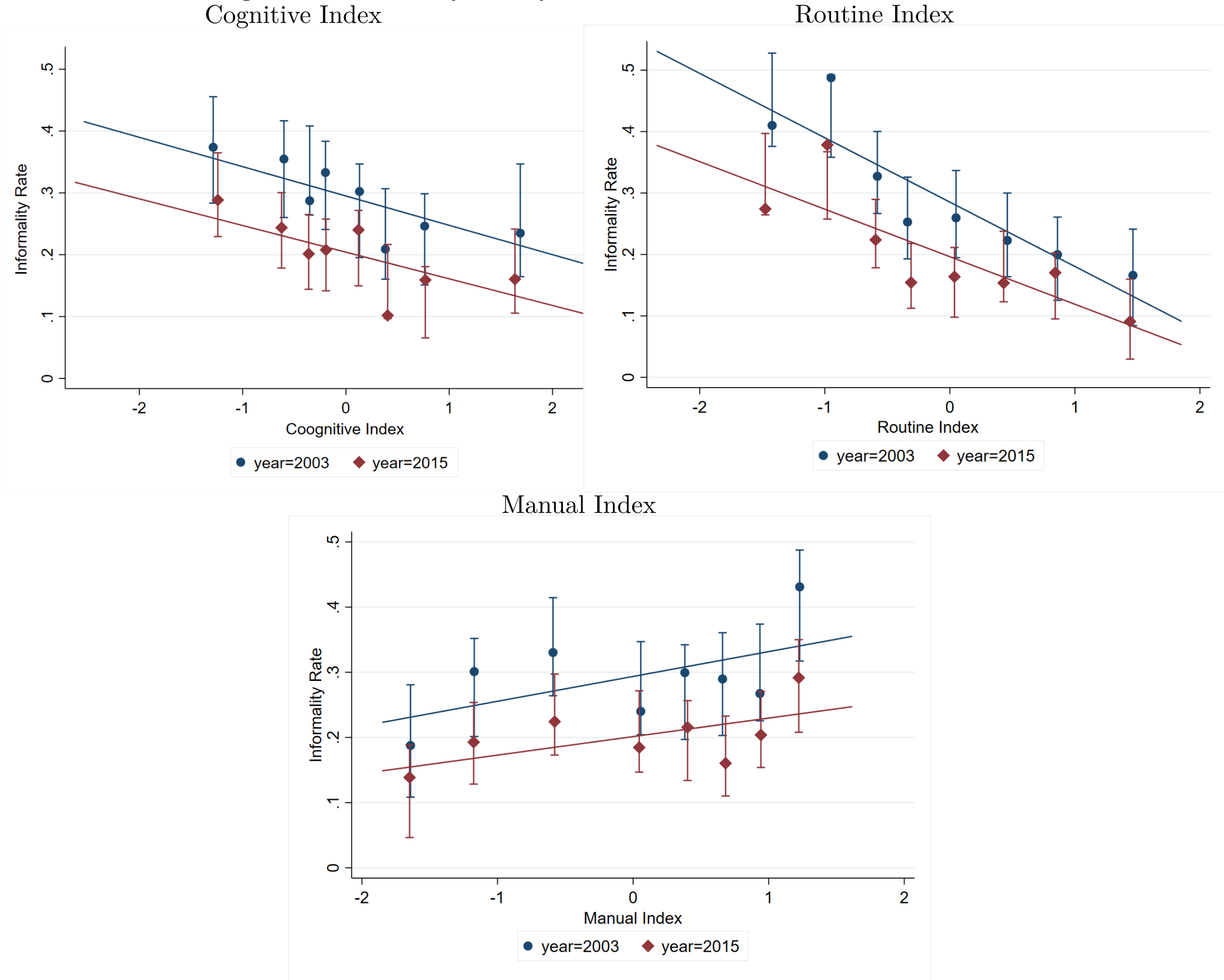
Source: Elaborated by the author from PNAD data, 25-65 year old employees from the private sector.

The formal labor market is predominantly composed of employees (87% in 2003 and 85% in 2014) whereas the informal market has a substantially higher number of self-employed individuals (47% and 52% against 7% and 10% in 2003 and 2015, respectively). The comparison between these two years shows a structural composition difference between both sectors (steady in time). This peculiarity of the informal sector is one potential source of differences in the likelihood of being informal and can potentially help explain divergences in the polarization context.

In Figure 8, we analyze the informality rate for each task level. We can see that it is decreasing in the level of requirement for cognitive and routine tasks. On the other hand, the informality rate is increasing for those occupations which require manual tasks. The evidence is in line with the descriptive statistics of Table 9.

We analyze the correlation of task indexes and informality rate in binscatter plots to informally test the linearity and monotonicity of the regression function. At first glance, it seems that the informality rate fell between 2003 and 2015 in all cases (all years and all levels of each task). However, the analysis of the confidence interval of bins suggests that the informality rate dropped only (statistically significant) for positive levels of manual requirement. In occupations intensive in cognitive or routine tasks, the difference seems to be in the confidence interval. Yet, the informality rate seems to have decreased significantly in occupations less intensive in the routine task.

Figure 8 – Informality rate by tasks in binned scatter - 2003-2015



Source: Elaborated by the author from the 2003-2015 PNAD dataset. Notes: The figure plots binned scatter between the informality rate and task indexes for the years 2003 and 2015 (number of bins: 8; polynomial of degree 1 with one smoothness constraints used for constructing confidence intervals.)

The previous analysis allow us to conjecture there was a migration to the formal sector of workers who performed a vast combination of the tree tasks, but specially those intensive in manual tasks. Thus, the increase in manual employment occured in the formal sector may not only be due to the leaving of workers who performed routine tasks – those who were replaced by computer and machines as technological changes happened –, as the polarization models propose. It may also be due to the incorporation of labor that were previously informal.

4.4 EMPIRICAL STRATEGY

This section presents the empirical strategy on estimating the impact of cognitive, routine, and manual tasks in the probability of an individual being informal and in the wage gap between these two sectors. First, the section exhibits a three-stage model for estimating the role of tasks on the likelihood of being informal using PNAD data to determine earnings and selection between formal and informal sectors and then estimate structural probit. The section also presents the estimations of Oaxaca-Blinder decomposition to investigate the role of tasks in the wage gap between formal and informal.

4.4.1 Modeling the Propensity of Being Informal

Based on the evidence above, we simultaneously model the participation decision and earnings of being informal. We assume the selection of workers into sectors does not occur randomly, so the wage differences do not have a clear causal relation. In this case, individuals rationally choose to be allocated in the formal or informal sector based on their utility derived from each sector. Our hypothesis is based on the endogenous segmentation theory of Meghir, Narita and Robin (2015) and takes the firm's decision as given. We are interested in estimating the role of tasks required by occupations when they make their decision about which sector they will work.

In our model, the probability of being informal depends on the utility gained in each sector. The likelihood is assumed to be a linear function of individuals' earnings and characteristics, as in Carneiro and Henley (2001). We also include our parameters of interest θ_j, where $j = \{c, r, m\}$ represents cognitive (c), routine (r) and manual (m) tasks. Formally, we have:

$$Pr(U_i^I - U_i^F) = Pr(\alpha + \beta(ln(Y_i^I) - ln(Y_i^F)) + \nu_1\theta_{ci} + \nu_2\theta_{ri} + \nu_3\theta_{mi} + \gamma X_i + \epsilon_i \geq 0) \quad (4.1)$$

where U_i^I is the utility of choosing the informal sector while U_i^F is the same for the formal sector. In addition, Y_i^I represents the earnings received if the individual is in the informal sector, while Y_i^F is the same for formal one. As established above, θ_j with $j = \{c, r, m\}$ represents the task required by occupation in which the individual is located. Finally, X_i is the vector of individual characteristics such as age, experience, education, etc. As usual, ϵ_i is an idiosyncratic error, α is the intercept and β, ν and γ are the coefficient vectors.

In this model, we do not observe both Y_i^I and Y_i^F at the same time because individuals are only in one of the two states – one of the two sectors indeed. We perform a Heckman (1979) two-stage procedure to estimate earnings in both sectors, such that we can estimate Equation (4.1) in the third step through a probit regression. This procedure is possible under the assumption that $\epsilon_i \sim N(0, \sigma^2)$.

Let S be a binary indicator such that $S_i = 1$ if $U_i^I - U_i^F > 0$ and $S_i = 0$, otherwise. In the first step, we estimate the selection mechanism probit (equation (4.2)) by maximum likelihood to obtain estimates of φ:

$$Pr(S_i) = Pr(\varphi W_i + u_i \geq 0) \quad (4.2)$$

where $Pr(S_i = 1) = \Phi(\varphi W_i)$ and $Pr(S_i = 0) = 1 - \Phi(\varphi W_i)$. The vector W_i includes all variables of X_i, our parameters of interest θ_{ci}, θ_{ri} and θ_{mi}, and a new vector of variables Z_i, which includes variables such as dummy to whether individuals are head of household, or if she receives other income besides the labor wages and if she has multiple jobs. For each observation in the selected sample, we compute the selectivity correction term (inverse Mills ratio) $\hat{\lambda} = \phi(\hat{\varphi}W_i)/\Phi(\hat{\varphi}W_i)$ and $\hat{\delta} = \hat{\lambda}(\hat{\lambda} + \hat{\varphi}W_i)$.

We assume that earnings are given by semi-log functions such as:

$$\begin{aligned} \ln Y_i^I &= \psi_1^I \theta_{ci}^I + \psi_2^I \theta_{ri}^I + \psi_3^I \theta_{mi}^I + \mu^I X_i^I + \varepsilon_i^I \\ \ln Y_i^F &= \psi_1^F \theta_{ci}^F + \psi_2^F \theta_{ri}^F + \psi_3^F \theta_{mi}^F + \mu^F X_i^F + \varepsilon_i^F \end{aligned} \tag{4.3}$$

with $(u_i, \varepsilon_i) \sim$ bivariate normal$[0, 0, 1, \sigma_\varepsilon, \rho]$.

In the second stage, we estimate equations (4.3) by incorporating inverse Mills ratio $(\hat{\lambda})$:

$$\begin{aligned} \mathrm{E}[ln(Y_i^I)/S_i = 1] &= \psi_1^I \theta_{ci}^I + \psi_2^I \theta_{ri}^I + \psi_3^I \theta_{mi}^I + \mu^I X_i^I + \rho\sigma_\epsilon \hat{\lambda^I}(\hat{\varphi}W_i) \\ \mathrm{E}[ln(Y_i^F)/S_i = 0] &= \psi_1^I \theta_{ci}^I + \psi_2^I \theta_{ri}^I + \psi_3^I \theta_{mi}^I + \mu^I X_i^I + \rho\sigma_\epsilon \hat{\lambda^F}(\hat{\varphi}W_i) \end{aligned} \tag{4.4}$$

The results from equation (4.4) are used to construct predicted earnings in each state for all individuals. Such predictions are then used to replace Y^I and Y^F in the probit estimation of equation (4.1).

4.4.2 Oaxaca-Blinder Decomposition of Formal-Informal Wage Gap

As seen in Figure 7, the wage dynamics did not follow the polarization pattern observed in other countries, namely an increase in the lower and upper tail of wage distributions (occupation's median wage percentile ranking), as documented by the literature. In chapter 2 of this thesis, we saw the importance of analyzing task prices rather than only looking at the mean or median wage of occupations. In this paper, we analyze the specific task role in the formal-informal wage gap once wages are important elements in the decision of agents in their allocation into each sector.

Over the period analyzed, there was an important change in the Brazilian labor market. The formal-informal wage gap (favoring the formal sector), measured by the log of hourly wages, decreased from 31% in 2003 to 17% in 2015. We are interested in understanding how the movement in the market for each task can explain this reduction.

We adapt the Oaxaca-Blinder decomposition technique (OAXACA, 1973; BLINDER, 1973) to the framework of our interest. This technique has been refined over time, with contribution of Reimers (1983), Cotton (1988), Neumark (1988) Fortin (2008) and Stoczynski (2015). In particular, we set up our model in the following way. Individuals work in one of two sectors: formal ($s_i = 1$) or informal ($s_i = 0$). Their wages are:

$$\begin{aligned} \ln\,(Y_i^s) &= K_i\beta_1 + \varepsilon_{1i},\ \text{ if } s_i = 1 \\ \ln\,(Y_i^s) &= K_i\beta_0 + \varepsilon_{0i},\ \text{ if } s_i = 0, \end{aligned} \tag{4.5}$$

where ln (Y_i^s) represents logarithm earnings in sector $s \in \{1, 0\}$ and K_i is the vector composed of the parameters of our interest (θ_j) and the individual characteristics of the workers (X_i), and $E[\varepsilon_{1i}|K_i, si] = E[\varepsilon_{0i}|K_i, si] = 0$.

The formal-informal wage gap is therefore given by:

$$\begin{aligned} E[\ln(Y_i^s)|si = 1] - E[\ln(Y_i^s)|si = 0] = \quad & (E[Ki|si = 1] - E[Ki|si = 0])\beta_0 \\ & + E[Ki|si = 1](\beta_1 - \beta_0). \end{aligned} \tag{4.6}$$

The first term on the right-hand side of equation (4.6) $((E[Ki|si = 1] - E[Ki|si = 0])\beta_0)$ represents the difference in wages due to the observable characteristics of individuals in both sectors, which literature calls *explained component of the gap*. The second term $(E[Ki|si = 1](\beta_1 - \beta_0))$ refers to characteristics we do not observe, as abilities of workers, for example. We obtain these unobserved components through the difference of coefficients obtained in each sector applied to the characteristics of individuals who compose the formal market. This last component is called *unexplained term of the decomposition*.

Similarly, we can alternate the coefficients in equation (4.6) to obtain:

$$\begin{aligned} E[\ln(Y_i^s)|si = 1] - E[\ln(Y_i^s)|si = 0] = \quad & (E[Ki|si = 1] - E[Ki|si = 0])\beta_1 \\ & + E[Ki|si = 0](\beta_1 - \beta_0) \end{aligned} \tag{4.7}$$

The difference between equations (4.6) and (4.7) is in the group chosen to be the basis, that is, which the structure wage is assumed. In the first equation, the base is the characteristics of individuals who compose the formal market, whereas the base in the second equation is the characteristics of individuals who compose the informal market.

There is no consensus in the literature about which wage structure (the group of comparison) is better to be the base (either equation (4.6) or equation (4.7)). The choice of one of them may generate bias on the estimation. “The standard response has been to suggest alternative wage structures to solve this comparison group choice problem. Such an approach is referred to as *generalized Oaxaca–Blinder*” (STOCZYNSKI, 2015).

The alternative decomposition is given by:

$$\begin{aligned} E[\ln(Y_i^s)|si = 1] - E[\ln(Y_i^s)|si = 0] = \quad & (E[Ki|si = 1] - E[Ki|si = 0])\beta^* \\ & + E[Ki|si = 1](\beta_1 - \beta^*) - E[Ki|si = 0](\beta^* - \beta_0) \end{aligned} \tag{4.8}$$

where β^* is the set of comparison coefficients. Such an alteernative is often seen as the “competitive” wage structure. Again, the first term of equation (4.8) $(E[Ki|si =$

$1]-E[Ki|si=0])\beta^*)$ is the *explained component of the gap* and the second term $(E[Ki|si=1](\beta_1-\beta^*)-E[Ki|si=0](\beta^*-\beta_0))$ is the *unexplained component of the gap.*

In this version of Oaxaca-Blinder decomposition, some authors as Cotton (1988) interpret the *unexplained component* as the amount by which productivity characteristics of advantaged workers (in our case, formal workers) are overvalued $(E[Ki|si=1](\beta_1-\beta^*))$ plus the amount by which the productivity characteristics of disadvantaged workers (informal workers) are undervalued $E[Ki|si=0](\beta^*-\beta_0)$. In this paper, however, we are only interested in the importance of tasks in the *explained component of the wage gap ($E[Ki|si=1]-E[Ki|si=0])\beta^*$).*

Many proposals of alternative sets of comparison coefficients (β^*) have been made by literature. They usually refer to the linear combination of β_0 and β_1: $\beta^*=\lambda\beta_0+(1-\lambda)\beta_1$, where $\lambda\in[0,1]$. The divergence are often about the definition of the weighting factor λ. Reimers (1983), for example, suggested $=1/2$ while Cotton (1988) proposed $\lambda=P[s_i=1]$ (the proportion of advantaged workers). We follow Stoczynski (2015), who suggested $\lambda=P[s_i=0]\beta_1+P[s_i=1]\beta_0$. The idea is to use the population proportion of one group (i.e. advantaged workers) as a counterfactual for the other group (disavantaged workers).

4.5 RESULTS

This section presents the estimation's results of the propensity of being informal and the Oaxaca-Blinder decomposition of formal-informal wage gap.

4.5.1 Propensity of Being Informal

In this section, we present the main results of our first empirical exercise, namely the estimation of the impact of tasks on the propensity of being informal. Table 10 reports the results for the reduced form selection equation (equation (4.2)) and for selective corrected estimates of earnings functions (equation (4.3)). The tree estimations include the tasks of our interest, a dummy variable for gender (women=1), race (nonwhite=1), and urban area (yes=1). Also, we add age, education (year of schooling), tenure, and two dummies for self-employed (yes=1) and employer (yes=1) to our model. The selection equation also includes a dummy variable for the position in the household, multiple job-holding, and other household income for identification purposes.

As one can see below, coefficients are strongly statistically significant in the selection probit for both years. The influence of each component will be discussed later, in the estimation of structural probit. The results for the estimation of earnings functions, in turn, show that cognitive is the most relevant task to determine wages. It have a positive effect on wages in both sectors and both years, although its importance had slightly reduced between 2003 and 2015 in the informal one. Its coefficients appear to be greater than the returns of education, which have traditionally been seen as the principal component of wages (Mincer).

It is also possible to see that routine tasks contribute negatively to wages, and their coefficient became more negative in 2015 in the informal sector. The return of manual tasks, which were negative in both sectors in 2003, turned to positive in 2015. From these results, we can conjecture the following: (i) the demand for routine tasks, which was low in the informal sector, may have declined even more over the years, reducing the pressure on wages; (ii) the demand for manual tasks may have grown in both sectors between 2003 and 2015, putting pressure on wages. The formal sector coefficients are in line with the results found in chapter 2 of this thesis.

The coefficients of the covariate variables are in line with what was expected. Being female or individuals nonwhite decreases earnings in both sectors. In the formal sector, women's returns became less negative between 2003 and 2015, which did not occur to nonwhites. On the other hand, age, education, and the fact that individuals are self-employed or employer increase earnings. Education, in particular, had its importance reduced over time in both sectors, possibly due to the growing importance of cognitive tasks, independently of the education level. Finally, notice that we also control the earnings functions for the event of individuals being self-employed or employer. The categories vary between sectors according to the contribution to social security in our definition. Nevertheless, they explain an essential part of earnings, notably in the informal sector.

The estimate for the selectivity correction term is given by coefficient "lambda" (λ) in the earnings functions. In the informal sector, λ has a positive signal and is statistically significant, showing that the workers in this sector have a comparative advantage of working in it. This is also true for the formal sector in 2015, but the result no longer holds in 2003, when the coefficient was not statistically different from zero. It is remarkable that the degree of correlation between the error terms in the selection equation and the earnings equation (ρ) show a positive signal for all, except the formal sector in 2003, which corroborates the hypothesis of selection bias[8].

For the structural probit regression (equation (4.1), we use the tree tasks, dummy for gender, race, urban, head of household multiple job holding, other household income, self-employed and employer. The results are reported in Table 11. The coefficients of the predicted wage differential are positive and strongly significant. Observe that one percentage point increase in the difference between the log earnings raises the probability of being informal in 1.78 percentage points (p.p.) in 2003 and 1.68 (p.p.) in 2015. One can also see that, among the tasks required by occupations, cognitive had a negative impact on the likelihood of individuals working in the informal sector in 2003, and no significant

[8] A condition to λ be zero, what means no selection bias, is $\rho = 0$, since $\sigma_u > 0$

Table 10 – Selectivity corrected estimates of earnings functions for formal and informal employment

	Selection Probit - Pr(Informal)		Formal Earnings (Y^F)		Informal Earnings (Y^I)	
	2003	2015	2003	2015	2003	2015
COGNITIVE	-0.01	-0.05	0.15	0.16	0.15	0.12
	(0.0066)***	(0.0065)***	(0.0032)***	(0.0025)***	(0.0056)***	(0.0056)***
Routine	-0.23	-0.25	-0.03	-0.01	-0.05	-0.11
	(0.0062)***	(0.0064)***	(0.0061)***	(0.0044)	(0.0088)***	(0.0088)***
Manual	-0.03	0.02	-0.01	0.01	-0.02	0.01
	(0.0068)	(0.0068)***	(0.0036)***	(0.0025)**	(0.0054)***	(0.0054)*
Woman	-0.06	-0.05	-0.31	-0.26	-0.39	-0.38
	(0.0133)	(0.0123)***	(0.0057)***	(0.0044)***	(0.0107)***	(0.0107)***
Nonwhite	0.22	0.19	-0.18	-0.15	-0.15	-0.10
	(0.0110)***	(0.0109)***	(0.0074)***	(0.0047)***	(0.0112)***	(0.0112)***
Age	-0,01	-0.01	0.02	0.01	0.01	0.00
	(0.0006)***	(0.0005)***	(0.0004)***	(0.0002)***	(0.0006)***	(0.0006)***
Urban	-0.15	-0.27	0.05	0.05	0.09	0.12
	(0.0267)***	(0.0223)***	(0.0148)***	(0.0112)***	(0.0206)***	(0.0206)***
Education	-0.07	-0.07	0.07	0.06	0.05	0.02
	(0.0016)***	(0.0016)***	(0.0017)***	(0.0012)***	(0.0024)***	(0.0024)***
Tenure	0.00	0.00	0.00	0.00	0.00	0.00
	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)**	(0.0000)***
Self-employed	1.57	1.40	0.24	0.06	0.59	0.67
	(0.0148)***	(0.0131)***	(0.0479)***	(0.0313)**	(0.0459)***	(0.0459)***
Employer	0.64	0.51	0.60	0.43	0.88	0.81
	(0.0238)***	(0.0256)***	(0.0199)***	(0.0132)***	(0.0287)***	(0.0287)***
Head of Household	-0.13	-0.06				
	(0.0127)***	(0.0112)***				
More than 1 job held	0.06	0.08				
	(0.0328)***	(0.0394)*				
Other Income	0.01	-0.01				
	(0.0229)	(0.0328)				
Constant	0.61	0.29	5.10	6.29	4.61	5.65
	(0.0365)***	(0.0346)***	(0.0533)***	(0.0379)***	(0.0406)***	(0.0406)***
λ			-0.02	0.13	0.60	0.70
			(0.0488)	(0.0414)***	(0.0517)***	(0.0517)***
LR $\chi^2(14)$	22,049.52	22,379.06				
	(0.000)***	(0.000)***				
ρ			-0.03	0.27	0.74	0.84
σ			0.53	0.48	0.81	0.83
Obs.	70,813	81,634				

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***)p<0.01. The results report the regression of the first and second Heckman's steps defined by equations 4.2 and 4.4. See the text for details.

coefficient in 2015. Moreover, routine tasks also impact the probability of being informal negatively, whereas manual ones have a positive influence on being informal, both in 2003 and 2015.

The results also show that the coefficients of covariates are as expected. Women and nonwhite are significantly more likely to be employed in the informal sector. The coefficient of women increased between the two periods analyzed while the coefficient of nonwhite decreased at the same time. If an individual lives in an urban area or is head of household, the probability of being an informal decrease, and both coefficients are strongly significant. On the other hand, the “other income" coefficient was not significantly different from zero, which means the fact of having another income source does not influence the

Table 11 – Structural probit estimates for informal versus formal status

	2003	2015
Predicted $\ln(Y^I) - ln(Y^F)$	1.78	1.68
	(0.04085)***	(0.04217)***
Cognitive	-0.04	0.00
	(0.00722)***	(0.00737)
Routine	-0.20	-0.09
	(0.00729)***	(0.00869)***
Manual	0.01	0.01
	(0.00746)*	(0.00747)*
Woman	0.09	0.15
	(0.01565)***	(0.01495)***
Not_white	0.20	0.08
	(0.01225)***	(0.01251)***
Urban	-0.27	-0.38
	(0.02801)***	(0.02309)***
Head of Household	-0.09	-0.06
	(0.01369)***	(0.01222)***
More than 1 job held	0.07	0.06
	(0.03693)*	(0.04276)
Other Income	-0.01	0.02
	(0.02616)	(0.03754)
Self employment	0.99	0.36
	(0.01756)***	(0.02542)***
Employer	0.10	-0.14
	(0.02606)***	(0.02989)***
Constant	1.34	1.35
	(0.05007)***	(0.05814)***

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***) p<0.01.
The results report the structural probit regression (or third step) defined by equation 4.1. See the text for details.

likelihood of being informal. Dummies for self-employed and employer were considered in the estimation as controllers. Both coefficients are positive, except employers in the 2015 estimation, which became statistically negative.

In sum, the findings provided by the structural probit estimation are in line with the evidence we saw in section 4.3. In fact, even after controlling for selection bias and observables characteristics, the probability of being informal is negatively correlated to cognitive and routine and only positively to manual tasks.

4.5.2 Oaxaca-Blinder Decomposition

In this section, we present the results from Oaxaca-Binder decomposition, which allows us to understand the impact of tasks on the wage gap between formal and informal sectors. By doing so, we can control for various observable individual characteristics as age, experience, education, and others. The results are presented in Table 12.

We estimate two regressions: one analyzing only the impact of the three tasks on the formal-informal wage gap and another taking into account several other individual characteristics. In both types of regressions and both years analyzed, cognitive requirements account for a positive effect on the wage gap, whereas routine and manual requirements have little or no effect on it.

Considering the full regression, individuals' observable characteristics could explain 48.2 % of all this difference in 2003, with 51.8 % not explained by them. In 2015, a smaller proportion (39.6 %) could be explained by these same observable variables.

Among all the explanatory components, cognitive tasks' requirements are the second largest contributor to the wage gap in both years (the total impact of age – in linear and quadratic form – is less than 3 basis point). In 2003, it explained 9.6% of all difference between informal and formal wages (in log), while manual task contributed fewer to the wage gap (1.7%). However, the routine requirements were not significantly different from zero in de 2003 regression, appearing not contributing to the wage gap. In 2015, cognitive tasks' contribution to the gap jumped to 13%, becoming the one with the highest growth among the explanatory components of the wage gap between formal and informal. Furthermore, both routine and manual tasks did not affect the differential in that year.

The behavior of the observable covariate variables is in line with expectations. Education is the principal component explaining the formal-informal wage gap. According to the traditional Mincerian wage equations, years of schooling have a positive effect on wages and, as seen in descriptive statistics (tabela 9), they are higher, on average, among individuals in the formal sector. Therefore, the sign and magnitude of the coefficient associated with it as well as its importance (p-value) to explain the gap are not surprising. Observe also that both tenure and the wages of women contribute to narrow down the difference between formal-informal earnings. Finally, the percentage of nonwhite is a factor that increases the gap in both years.

Our findings show that cognitive tasks have an important role in the formal-informal wage gap. Over the years analized in this paper, their relevance in explaining the gap has increased from 9.6% to 13.0%, contributing to keeping the difference in favor of the formal sector (or, at least, to make the decline in this difference to be slower over time). The cognitive valorization in this sector analyzed in chapter 2 of this thesis may have worked to this happen. Manual and routine tasks have a limited effect on the wage gap in both years, such that both a low power to explain the closing of such a gap.

Table 12 – Oaxaca-Blinder decomposition of formal-informal wage gap

	2003				2015			
	Short		Full		Short		Full	
	Coef.	% of Diff.	Coef.	% of Diff.	Coef.	% of Diff.	Coef.	% of Diff.
Formal	0.96		0.96		2.09		2.09	
	(0.0042)***		(0.0042)***		(0.0026)***		(0.0027)***	
Informal	0.66		0.66		1.73		1.73	
	(0.0049)***		(0.0049)***		(0.0049)***		(0.0049)***	
Difference	0.30	100.0	0.30	100.0	0.36	100.0	0.36	100.0
	(0.0065)***		(0.0065)***		(0.0056)***		(0.0056)***	
Explained	0.07	23.8	0.14	48.2	0.08	23.0	0.14	39.6
	(0.0035)***		(0.0045)***		(0.0040)***		(0.0047)***	
Cognitive	0.07	24.0	0.03	9.6	0.072	19.8	0.05	13.0
	(0.0024)***		(0.0015)***		(0.0022)***		(0.0016)***	
Routine	0.00	1.1	0.00	-0.3	0.003	0.9	0.00	-0.3
	(0.0017)*		(0.0016)		(0.0023)		(0.0020)	
Manual	0.00	0.4	0.00	1.7	0.001	0.3	0.00	0.2
	(0.0010)***		(0.0009)***		(0.0014)		(0.0015)	
Woman			-0.03	-9.5			-0.04	-9.7
			(0.0011)***				(0.0014)***	
Nonwhite			0.02	7.3			0.02	6.1
			(0.0010)***				(0.0011)***	
Age			-0.10	-34.3			-0.07	-20.1
			(0.0064)***				(0.0050)***	
Age_sq			0.09	28.6			0.07	19.2
			(0.0051)***				(0.0045)***	
Metrop.			0.00	1.4			0.01	1.5
			(0.0004)***				(0.0006)***	
Tenure			-0.01	-2.3			-0.01	-3.4
			(0.0009)***				(0.0008)***	
Age start work			0.00	0.0			0.00	0.2
			(0.0007)				(0.0009)	
Years of school.			0.14	46.5			0.11	31.7
			(0.0031)***				(0.0025)***	
Unexplained	0.23	76.2	0.155	51.8	0.22	60.4	0.22	60.4
	(0.0067)***		(0.0060)***		(0.0064)***		(0.0060)***	
Observations	68,424		67,787		82,630		82,400	
Formal	43,511		43,081		59,552		59,383	
Informal	24,913		24,706		23,078		23,017	

Notes: Standard errors in parentheses.(*) p<0.10, (**) p<0.05, (***)p<0.01. The results report the regression of Oaxaca-Blinder Decomposition defined by equation 4.8. See the text for details.

4.6 CONCLUSION

This paper presents an estimation of the impact of cognitive, routine, and manual tasks in the probability of an individual being informal and in the wage gap between these two sectors. In order to do that, we employ a three-stage model for estimating the role of tasks on the likelihood of being informal using PNAD data to simultaneous determine earnings and selection between formal and informal sectors and then estimate structural probit. The results shows that only manual task has a positive impact on the probability of being informal. These results are in line with the evidence discussed in section 4.3. Even after controlling for selection bias and observables characteristics, the probability of being informal is negatively correlated to cognitive and routine while the only positive effect among the three tasks comes from manual ones.

We also estimated an Oaxaca-Blinder decomposition to investigate the role of tasks in the wage gap between formal and informal. The results show that cognitive is the most important task to explain the gap both in 2003 and 2015. It responds to 13% of the difference, which represents an increase in comparison with 2003 (9.6%), contributing to the maintenance of wage gap over time. Manual and routine tasks, on the other hand, have little effect on it, so that we can conclude that tasks show a low power to influence on the closing formal-informal wage gap.

This paper opens several paths for future research. For example, one can test the hypothesis, based on the evidence shown in this paper, that the increase in employment in the lower tail of the wage distribution of the formal sector (figure 7) of the economy actually came from the formalization process in Brazil. Individuals who perform manual tasks were incorporated into the formal sector in part due to technological changes, but also as a function of structural change in the labor market composition faced by many developing countries – the so-called formalization process. The restriction imposed by PNAD database, which do not follow the individuals over time, prevent us from performing this test. Nevertheless, future research may explore other database to overcome such a difficulty.

REFERENCES

ACEMOGLU, D.; AUTOR, D. Skills, tasks and technologies: Implications for employment and earnings. In: CARD, D.; ASHENFELTER, O. (Ed.). *Handbook of Labor Economics, Volume 4B*. North Holland: Elsevier, 2011. p. 1043–1171. 26, 43, 52, 54, 56, 58, 59, 60

ADAMCZYK, W.; FRANCA, M. T. A.; FOCHEZATTO, A. Technological unemployment in brazil: effects of automation on job market occupations. *XXII ENCONTRO DE ECONOMIA DA REGIÃO SUL, 2019*, 2019. 29

ATALAY, E. et al. The evolving u.s. occupational structure. *Working Paper*, 2018. 28, 37

AUTOR, D.; KATZ, L.; KRUEGER, A. Computing inequality: have computers changed the labor market? *Quarterly Journal of Economics*, v. 113I, p. 1169–1214, 1998. 55

AUTOR, D. H. Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, v. 29, n. 3, p. 3–30, 2015. 27

AUTOR, D. H.; DORN, D. The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, v. 103, n. 5, p. 1553–1597, 2013. 26, 41, 55

AUTOR, D. H.; KATZ, L. F.; KEARNEY, M. S. The polarization of the u.s. labor market. *American Economic Review*, v. 96, n. 2, p. 189–194, 2006. 26, 41, 51, 55

AUTOR, D. H.; LEVY, F.; MURNAME, R. J. The skill content of recent technological change: An empirical exploration. *The Quartely Journal of Economics*, v. 118, n. 4, p. 1279–1333, 2003. 19, 25, 39, 40, 42, 51, 55

BACOLOD, M. P.; BLUM, B. S. Two sides of the same coin: U.s. "residual" inequality and the gender gap. *Journal of Human Resources*, v. 45, n. 1, p. 197–242, 2010. 27, 41

BARON-COEN, S.; KNICKMEYER, R. C.; BELMONTE, M. K. Sex differences in the brain: Implications for explaining autism. *Science*, v. 310, p. 819–823, 2005. 28, 38

BEAUDRY, P.; GREEN, D. A.; SAND, B. M. The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, v. 34, n. S1, 2016. 26

BECKER, G. *The Economics of Discrimination*. Chicago: University of Chicago Press, 1974. 35

BLACK, S. E.; SPITZ-ONER, A. Explaining women's success: Technological change and skill content of women's work. *The Review of Economics and Statistics*, v. 92, n. 1, p. 187–194, 2010. 27

BLINDER, A. Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, n. 4, p. 436–455, 1973. 31, 83

BLINDER, A. How many us jobs might be offshorable? *Princeton University CEPS Working Paper*, n. 142, 2007. 56

BLINDER, A.; KRUEGER, A. B. Measuring offshorability: a survey approach. *Princeton University Working Paper*, 2008. 56

BORGHANS, L.; WEEL, B. ter; WEINBERG, B. A. People skills and the labor-market outcomes of underrepresented groups. *Industrial Labor Relations Review*, v. 67, n. 2, p. 287–334, 2014. 24, 27, 37

BOTELHO, F.; PONCZEK, V. Segmentation in the brazilian labor market. *Economic Development and Cultural Change*, v. 59, n. 231, p. 437 – 463, 2011. 66

BRYNJOLFSSON, E.; MITCHELL, T.; ROCK, D. What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, n. 108, p. 43–47, 2018. 26

BULLA, J. F. A. Essays on wage inequality in developing countries. *PhD thesis, Universitat Autônoma de Barcelona*, 2014. 52, 56, 58

CARNEIRO, F. G.; HENLEY, A. Modelling formal vs. informal employment and earnings: Micro-econometric evidence for brazil. *U of Wales at Aberystwyth Management Business Working Paper*, v. 2001, n. 15, 2001. 77, 82

CAVAGLIA, C.; ETHERIDGE, B. Job polarization, task prices and the distribution of task returns. *ISER Working Paper Series*, n. 9, 2017. 54, 59, 66

CHAPMAN, E. et al. Fetal testosterone and empathy: Evidence from the empathy quotient (eq) and the “reading the mind in the eyes” test. *Social Neuroscience*, v. 1, n. 2, p. 135–148, 2010. 24, 28, 38

CORSEUIL, C. H. L.; POOLE, J. P.; ALMEIDA, R. K. The impact of digital technologies on worker tasks: Do labor policies matter? *IPEA Discussion Paper*, n. 234, 2018. 52, 57, 58

CORTES, G. M. Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics*, v. 34, n. 1, p. 63–105, 2016. 32, 33, 54, 59

CORTES, G. M.; JAIMOVICH, N.; SIU, H. E. The "end of men" and rise of women in the high-skilled labor market. *NBER Working Paper*, n. 24274, 2018. 24, 28, 29, 32, 34, 35, 37, 41, 42, 43, 45

CORTES, G. M.; JAIMOVICH, N.; SIU, H. E. The "end of men" and rise of women in the high-skilled labor market. *Working Paper*, 2020. 29

COTTON, J. On the decomposition of wage differentials. *The Review of Economics and Statistics*, v. 70, n. 2, p. 236–243, 1988. 83, 85

CURI, A. Z.; MENEZES-FILHO, N. A. O mercado de trabalho brasileiro é segmentado? alterações no perfil da informalidade e nos diferenciais de salários nas décadas de 1980 e 1990. *Estudos Econômicos*, v. 36, n. 4, p. 867–899, 2006. 77

DEMING, D. J. The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, v. 132, n. 4, p. 1593–1640, 2017. 24, 26, 27, 37, 38, 41, 42, 47

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. “occupational tasks and changes in the wage structure. *University of British Columbia Working Paper*, 2011. 56

FORTIN, N. M. The gender wage gap among young adults in the united states: The importance of money versus people. *The Journal of Human Resources*, v. 43, n. 4, p. 884–918, 2008. 83

FREIJE, S. El empleo informal en america latina y el caribe: causas, consecuencias y recomendaciones de política. *Banco Interamericano de Desarrollo, Primer Seminario Tecnico de Consulta Regional sobre Temas Laborales*, 2001. 77

FUNCHAL, B.; JUNIOR, J. S. Understanding demand for skylls after technological trade liberalization. *Fucape Working Papers*, v. 40, 2013. 52, 57

GILLIGAN, C. *In A Different Voice: Psychological Theory and Women's Development*. [S.l.]: Harvard University Press, 1993. 27, 38

GOLDIN, C.; KATZ, L. *The Race between Education and Technology*. Cambridge: Harvard University Press, 2008. 55

GOOS, M.; MANNING, A. Lousy and lovely jobs: The rising polarization of work in britain. *Review of Economics and Statistics*, v. 89, n. 2, p. 118–133, 2007. 26, 51, 55

GOOS, M.; MANNING, A.; SALOMONS, A. Job polarization in europe. *American Economic Review*, v. 99, n. 2, p. 55–63, 2009. 55, 56

GOOS, M.; MANNING, A.; SALOMONS, A. Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, v. 104, n. 8, p. 2509–2526, 2014. 26, 56

GROSSMAN, G.; ROSSI-HANSBERG, E. Trading tasks: a simple theory of offshoring. *American Economic Review*, v. 98, n. 5, p. 1978–1997, 2008. 56

HALL, J. A. Gender effects in decoding nonverbal cues. *Psychological Bulletin*, v. 85, n. 4, p. 845–857, 1978. 28, 38

HARDY, W.; KEISTER, R.; LEWANDOWSKI, P. Educational upgrading, structural change and the task composition of jobs in europe. *Economics Of Transition*, n. 26, 2018. 41

HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica*, v. 8, p. 153–161, 1979. 20, 82

HERDEIRO, R.; MENEZES-FILHO, N.; KOMATSU, B. Explicando a evolução dos salários relativos por grupos de qualificação no brasil. *Policy Paper*, n. 38, 2019. 52, 57, 58

IKENAGA, T.; KAMBAYASHI, R. Task polarization in the japanese labor market: Evidence of a long-term trend. *Industrial Relations*, v. 55, n. 2, p. 267–293, 2016. 26, 56

KAMBOUROV, G.; SIOW, A.; TURNER, L. Relationship skills in the labor and marriage markets. *Society for Economic Dynamics 2014 Meeting Papers*, n. 155, 2014. 27

KATZ, L.; MURPHY, K. Changes in relative wages: supply and demand factors. *Quarterly Journal of Economics*, CVII, p. 35–78, 1992. 55

KELLER, W.; UTAR, H. International trade and job polarization: Evidence at the worker-level. *NBER Working Paper*, v. 22315, 2016. 56

KRAMARZ, D. C. F.; LEMIEUX, T. Changes in the relative structure of wages and employment: A comparison of the united states, canada, and france. *The Canadian Journal of Economics*, v. 32, n. 4, p. 843–877, 1999. 55

MAGNAC, T. Segmented or competitive labor markets. *Econometrica*, v. 59, n. 1, p. 165–187, 1991. 76, 77

MALONEY, W. F.; MOLINA, C. Are automation and trade polarizing developing country labor markets, too? *WB Policy Research Working Paper*, n. 7922, 2016. 52, 57

MCCANN, R. J. et al. Becker meets ricardo: Multisector matching with communication and cognitive skills. *The Journal of Law, Economics, and Organization*, 2010. 27, 47

MEDINA, C.; POSSO, C. Technical change and polarization of the labor market: Evidence for brazil, colombia and mexico. *Borradores de Economia 614, Banco de la Republica de Colombia*, 2010. 52, 57

MEGHIR, C.; NARITA, R.; ROBIN, J.-M. Wages and informality in developing countries. *Review of Economics and Statistics*, v. 105, n. 4, p. 1509–46, 2015. 77, 82

MENEZES-FILHO, N. A.; MENDES, M.; ALMEIDA, E. S. de. O diferencial de salários formal-informal no brasil: segmentação ou viés de seleção? *Revista Brasileira de Economia*, v. 58, n. 2, 2003. 77, 78

MINISTéRIO DO TRABALHO E EMPREGO. *Manual de Orientação da Relação Anual de Informações Sociais (RAIS): ano base 2011*. Brasília, DF, 2012. 60 p. 68

MURPHY, W. C. R. K. M.; ROMER, P. M. Wages, skills and technology in the united states and canada. *NBER Working Paper*, n. 6638, 1998. 55

NERI, M. C. Decent work and the informal sector in brazil. *Ensaios Econômicos*, n. 461, 2002. 77

NEUMARK, D. Employers' discriminatory behavior and the estimation of wage discrimination. *The Journal of Human Resources*, v. 23, p. 279–295, 1988. 83

OAXACA, R. Male-female wage differentials in urban labor markets. *International Economic Review*, n. 14, p. 693–708, 1973. 31, 83

PONCZEK, V. Skill segmentation of informal employment. *Encontro Nacional de Economia - ANPEC*, 2007. 77

PORTES, A.; SCHAUFFLER, R. Competing perspectives on the latin american informal sector. *Population and Development Review*, v. 19, n. 1, p. 33–60, 1993. 76

REIJNDERS, L. S. M.; VRIES, G. J. de. Job polarization in advanced and emerging countries: The role of task relocation and technological change within global supply chains. *GGDC Research memoranda*, n. 167, 2017. 19, 56, 59

REIMERS, C. W. Labor market discrimination against hispanic and black men. *The Review of Economics and Statistics*, v. 65, p. 570–579, 1983. 83, 85

RIVA, F. L. R. *Computerization, occupational tasks and the labor market: evidence from a natural experiment in Brazil.* Dissertação (Dissertation (Masters)) — Escola de Economia de São Paulo, 2016. 52, 57

ROY, A. Some thoughts on the distribution of earnings. *Oxford economic papers*, v. 3, p. 135–146, 1973. 76

SPITZ-OENER, A. Technical change, job tasks and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, v. 24, n. 2, p. 235–270, 2006. 26, 51, 55

STOCZYNSKI, T. On the decomposition of wage differentials. *IZA Discussion Paper*, n. 9036, 2015. 83, 84, 85

SULZBACH, V. 2020. Disponível em: <https://sites.google.com/view/vanessasulzbach>. 43

TANNURI-PIANTO, M.; PIANTO, D. M. Informal employment in brazil – a choice at the top and segmentation at the bottom: A quantile regression approach. *Anais do XXIV Encontro Brasileiro de Econometria*, v. 2, n. 15, 2002. 77

TINBERGEN, J. Substitution of graduate by other labor. *Kyklos*, v. 27, n. 2, p. 217–226, 1974. 57

ULYSSEA, G. Informalidade no mercado de trabalho brasileiro: uma resenha da literatura. *Revista de Economia Política*, v. 26, n. 4 (104), p. 596–618, 2006. 78

ULYSSEA, G. Firms, informality, and development:theory and evidence from brazil. *American Economic Review*, v. 108, n. 8, p. 2015–2047, 2018. 76, 77

WEINBERGER, C. The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, v. 96, n. 4, p. 849–861, 2014. 27, 37, 47

WOOLLEY, A. W. et al. Evidence for a collective intelligence factor in the performance of human groups. *Science*, v. 330, n. 29, p. 686–688, 2010. 24, 27

WORLD BANK GROUP. *Gestão de Pessoas e Folha de Pagamentos no Setor Público Brasileiro: o Que Os Dados Dizem? (Vol. 2) (Portuguese)*. Washington, D.C, 2019. 96 p. 67

APPENDIX A – ADDITIONAL FIGURES AND TABLES OF CHAPTER 2

A.1 ADITIONAL TABLES OF SUBSECTION 2.5.1

Table 13 – Grouping occupations (CBO2002)

GROUP	CBO2002
Group I	1210
	1221
	1222
	1223
	1224
	1225
	1226
	1227
	1231
	1232
	1233
	1234
	1236
	1237
	1238
	1311
	1312
	1313
	1313
Group II	1412
	1413
Group III	1424
	1427
Group IV	2030
	2011
	3253
Group V	2141
	2629
Group VI	2143
	2021
	2122
Group VII	2145
	2222
	2211
Group VIII	2212
	3012
	3201
Group IX	2221
	2034
	2140
Group X	2233
	5193
Group XI	2311
	3311
Group XII	2313
	3312
	3321
Group XIII	2331
	3313
Group XIV	2332
	3322
Group XV	2344
	2033
Group XVI	2347
	2514
Group XVII	2394
	3331
	3341
Group XVIII	2512
	2525
Group XIX	2521
	2526
Group XX	2612
	4231
Group XXI	2613
	3712
Group XXII	2628
	3761
Group XXIII	3003
	3001
Group XXIV	3116
	3191
	7610
Group XXV	3133
	3135
Group XXVI	3141
	3142
Group XXVII	3182
	3186
Group XXVIII	3183
	3187
Group XXIX	3184
	3192
Group XXX	3211
	3212
Group XXXI	3225
	3226
Group XXXII	3412
	3413
Group XXXIII	3523
	2012
Group XXXIV	3532
	2532
Group XXXV	5132
	2711
	5136
Group XXXVI	5134
	5135
Group XXXVII	5163
	5102
Group XXXVIII	5173
	5103
	5153
Group XXXIX	7251
	7255
Group XL	7411
	7401
Group XLI	7421
	9152
Group XLII	7601
	7605
Group XLII	7614
	7654
Group XLIV	7621
	7620
Group XLV	7641
	7640
Group XLVI	7734
	7735
Group XLVII	7817
	7813
Group XVIII	7841
	7801
Group XLIX	8112
	8131
Group L	8117
	3114
Group LI	8417
	3250
Group LII	8421
	8486
Group LIII	9109
	9102
Group LIV	9113
	3144
Group LV	9501
	9502
	9503
Group LVI	9511
	9513

Source: Elaborated by the author.

Table 14 – Occupations that did not exist in 1994 or had few employees in 2017 or had no reference in the CBO2002 book

CBO 2002	Description
1130	Dirigentes de Povos Indígenas, de Quilombolas e Caiçaras
1141	Dirigentes de Partidos Políticos
1142	Dirigentes e Administradores de Entidades Patronais e dos Trabalhadores e de Outros Interesses Socieconômicos
1143	Dirigentes e Administradores de Entidades Religiosas
1144	Dirigentes e Administradores de Organizações da Sociedade Civil Sem Fins Lucrativos
2263	Profissionais das terapias criativas,equoterápicas e naturológicas
2514	Filósofos
2527	Sem referência
3519	Sem referência
5121	Trabalhadores dos Serviços Domésticos em Geral
5167	Astrólogos e Numerólogos
5168	Esotéricos e Paranormais
5198	Profissionais do Sexo
7911	Artesãos

Source: Elaborated by the author.

Table 15 – Occupations that was not referenced in CBO1994 book

CBO 1994		
2337	19930	28704
2425	19940	28720
5390	19950	28736
7420	20496	32784
7640	20512	32800
11090	20528	32816
16290	20544	32832
16400	24592	59975
16416	24608	59980
16432	24624	83405
16448	24640	92185
19390	28688	

Source: Elaborated by the author.

A.2 TOP AND BOTTOM TEN OCCUPATIONS OF EACH TASK

Table 16 reports the top and bottom ten occupations with the highest and lower level of Social Skills Intensity, respectively, in 2000 and 2017. Occupations of the top in 1994 are dominated by management positions (directors and managers), lawyers, social workers, and others (even economists). The picture of the top 10 occupations in social skills did not change too much until 2017. On the bottom side are the occupations that involve working with natural resources such as stones, ores, ceramics, and others when we look to the index in 2000. In 2017, occupations with more manual tasks took place in the bottom side, like laundry and leather tanning workers.

Table 17 shows the same to Cognitive Tasks Intensity. Engineers, directors, and managers were the principal occupations in the top 10 highest levels of Cognitive Skills. In 2017, engineers were more predominant in the top 10, including statisticians and some managers. In the other hand, models, laundry and gold miners and salt operators are in the bottom side in both years. The list in 2000 also includes, for example, funeral services workers and stone beneficiation workers. Meanwhile, workers in building management services, in telemarketing and leather tanning are in the list in 2017.

Concerning Routine Skills Intensity, occupations in Top 10 in 2000 were machine and equipment operators, ore processing, and leather tanning, as we can see in Table 18. In 2017, equipment operators' occupations were predominant, and cashiers and ticket agents became part of the list. Occupations involving programmers, evaluators, teaching counselors, lawyers, and veterinarians were the ones with less routine content in 2000.

Table 19 reports the top and down ten occupations according to the Manual Skills Intensity. Occupation on the top was dominated by police, firefighters, guards, athletes, and physical education professionals in 2000, while carpenters, prospectors, braiding of metal structures metal drawers took place in the top ten occupations in 2017. At the bottom of routine content intensity were archivists, lawyers, insurance, stock, and financial assets brokers, notaries, chemicals, accountants, among others in coth years.

Table 16 – Top and bottom 10 occupations in the social skill intensity Distribution - 2000 and 2017

Social Skills Intensity 2000		**Social Skills Intensity 2017**	
Top 10 **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**	**Top 10** **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**
Diretores Gerais	Trabalhadores de Beneficiamento de Pedras Ornamentais	Tabeliães e Registradores	Lavadores e Passadores de Roupa, à Mão
Advogados	Ceramistas (Preparação e Fabricação)	Assistentes Sociais e Economistas Domésticos	Supervisores da Indústria Têxtil
Gerentes de Suprimentos e Afins	Trabalhadores da Fabricação de Cerâmica Estrutural para Construção	Ministros de Culto, Missionários, Teólogos e Profissionais Assemelhados	Trabalhadores da Preparação do Curtimento de Couros e Peles
Gerente de Pesquisa e Desenvolvimento	Trabalhadores de Beneficiamento de Minérios	Advogados	Trabalhadores do Curtimento de Couros e Peles
Gerentes de Tecnologia da Informação	Afiadores e Polidores de Metais	Gerentes de Suprimentos e Afins	Supervisores na Indústria do Curtimento
Assistentes Sociais e Economistas Domésticos	Trabalhadores de Estruturas de Alvenaria	Gerentes de Comercialização, Marketing e Comunicação	Operadores de Usinagem Convencional de Madeira
Economistas	Trabalhadores na Operação de Máquinas de Concreto Usinado	Gerentes de Operações de Serviços em Instituição de Intermediação Financeira	Trabalhadores de Tecelagem Manual, Tricô, Crochê, Rendas e Afins
Gerentes de Produção e Operações em Empresa da Indústria Extrativa, de Transformação e de Serviços de Utilidade Pública	Operadores de Instalações e Equipamentos de Fabricação de Materiais de Construção	Gerentes de Recursos Humanos e de Relações do Trabalho	Trabalhadores de Moldagem de Metais e de Ligas Metálicas
Gerentes de Operações de Serviços em Empresa de Transporte, de Comunicação e de Logística	Operadores na Preparação de Massas para Abrasivo, Vidro, Cerâmica, Porcelana e Materiais de Construção	Gerentes de Operações Comerciais e de Assistência Técnica	Trabalhadores nos Serviços de Administração de Edifícios
Arquitetos e Urbanistas	Operadores de Equipamentos de Fabricação e Beneficiamento de Cristais, Vidros, Cerâmicas, Porcelanas, Fibras de Vidro, Abrasivos e Afins	Diretores Gerais	Trabalhadores da Preparação de Artefatos de Tecidos, Couros e Tapeçaria

Source: Elaborated by the author.

Table 17 – Top and bottom 10 occupations in the cognitive skill distribution - 2000 and 2017

Cognitive Skills 2000		**Cognitive Skills 2017**	
Top 10 **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**	**Top 10** **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**
Engenheiros Químicos e Afins	Modelos	Engenheiros Civis e Afins	Modelos
Diretores Gerais	Filólogos,Tradutores, Intérpretes e Afins	Engenheiros Químicos e Afins	Trabalhadores nos Serviços de Administração de Edifícios
Gerentes de Tecnologia da Informação	Alimentadores de Linhas de Produção	Engenheiros de Minas e Afins	Lavadores e Passadores de Roupa, à Mão
Gerentes de Pesquisa e Desenvolvimento e Afins	Lavadores e Passadores de Roupa, à Mão	Engenheiros Metalurgistas de Materiais e Afins	Trabalhadores da Preparação do Curtimento de Couros e Peles
Engenheiros Civis e Afins	Trabalhadores Auxiliares dos Serviços Funerários	Pesquisadores de Engenharia e Tecnologia	Supervisores na Indústria do Curtimento
Economistas	Trabalhadores Operacionais de Conservação de Vias Permanentes (Exceto Trilhos)	Engenheiros Mecânicos e Afins	Supervisores da Indústria Têxtil
Engenheiros Eletricistas, Eletrônicos e Afins	Trabalhadores dos Serviços Funerários	Profissionais de Estatística	Trabalhadores do Curtimento de Couros e Peles
Gerentes de Operações de Serviços em Empresa de Transporte, de Comunicação e de Logística (Armazenagem e Distribuição)	Garimpeiros e Operadores de Salinas	Gerentes Administrativos, Financeiros, de Riscos e Afins	Operadores de Telemarketing
Engenheiros Mecânicos e Afins	Trabalhadores de Beneficiamento de Pedras Ornamentais	Engenheiros de Produção, Qualidade, Segurança e Afins	Vendedores em Domicílio
Gerentes de Suprimentos e Afins	Trabalhadores na Operação de Máquinas de Terraplenagem e Fundações	Gerentes de Produção e Operações em Empresa da Indústria Extrativa, de Transformação e de Serviços de Utilidade Pública	Trabalhadores de Tecelagem Manual, Tricô, Crochê, Rendas e Afins

Source: Elaborated by the author.

Table 18 – Top and bottom 10 occupations in the routine skill distribution - 2000 and 2017

Routine Skills 2000		Routine Skills 2017	
Top 10 **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**	**Top 10** **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**
Trabalhadores na Operação de Máquinas de Concreto Usinado	Ministros de Culto, Missionários, Teólogos e Profissionais Assemelhados	Engenheiros Agrimensores e Engenheiros Cartógrafos	Modelos
Trabalhadores de Beneficiamento de Minérios	Advogados	Farmacêuticos	Tabeliães e Registradores
Preparadores e Operadores de Máquinas-Ferramenta Convencionais	Programadores, Avaliadores e Orientadores de Ensino	Operadores Polivalentes de Equipamentos em Indústrias Químicas, Petroquímicas e Afins	Professores de Nível Superior na Educação Infantil
Ajustadores Mecânicos Polivalentes	Veterinários e Zootecnistas	Supervisores de Produção em Indústrias Químicas, Petroquímicas e Afins	Ministros de Culto, Missionários, Teólogos e Profissionais Assemelhados
Operadores de Instalações e Equipamentos de Fabricação de Materiais de Construção	Gerentes de Produção e Operações em Empresa da Indústria Extrativa, de Transformação e de Serviços de Utilidade Pública	Operadores de Equipamentos de Destilação, Evaporação e Reação	Trabalhadores Polivalentes da Confecção de Artefatos de Tecidos e Couros
Operadores de Máquinas de Fabricar Papel e Papelão	Corretores de Imóveis	Operadores de Equipamentos de Coqueificação	Supervisores na Confecção de Calçados
Preparadores de Pasta para Fabricação de Papel	Trabalhadores dos Serviços Funerários	Operadores de Máquinas à Vapor e Utilidades	Trabalhadores Artesanais da Confecção de Calçados e Artefatos de Couros e Peles
Trabalhadores de Forjamento de Metais	Geólogos, Oceanógrafos, Geofísicos e Afins	Operadores de Equipamentos de Filtragem e Separação	Trabalhadores de Acabamento de Calçados
Trabalhadores do Curtimento de Couros e Peles	Técnicos de Vendas Especializadas	Operadores de Equipamentos de Moagem e Mistura de Materiais (Tratamentos Químicos e Afins)	Trabalhadores da Preparação da Confecção de Calçados
Trabalhadores da Preparação do Curtimento de Couros e Peles	Supervisores de Vendas e de Prestação de Serviços	Caixas e Bilheteiros (Exceto Caixa de Banco)	Trabalhadores da Classificação de Fibras Têxteis e Lavagem de Lã

Source: Elaborated by the author.

Table 19 – Top and bottom 10 occupations in the manual skill distribution - 2000 and 2017

Manual Skills 2000		**Manual Skills 2017**	
Top 10 **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**	**Top 10** **CBO 2002 - Descrição**	**Bottom 10** **CBO 2002 - Descrição**
Trabalhadores Subaquáticos	Arquivistas e Museólogos	Trabalhadores Auxiliares dos Serviços Funerários	Tabeliães e Registradores
Trabalhadores nos Serviços de Administração de Edifícios	Ministros de Culto, Missionários, Teólogos e Profissionais Assemelhados	Garimpeiros e Operadores de Salinas	Gerentes de Recursos Humanos e de Relações do Trabalho
Policiais, Guardas- Civis Municipais e Agentes de Trânsito	Tabeliães e Registradores	Carpinteiros Navais	Contadores e Afins
Encanadores e Instaladores de Tubulações	Advogados	Carpinteiros de Carrocerias e Carretas	Profissionais de Estatística
Trançadores e Laceiros de Cabos de Aço	Corretores de Seguros	Trabalhadores na Operação de Máquinas de Terraplenagem e Fundações	Corretores de Valores, Ativos Financeiros, Mercadorias e Derivativos
Árbitros Desportivos	Técnicos de Seguros e Afins	Trabalhadores Aquaviários	Vendedores em Domicílio
Profissionais da Educação Física	Leiloeiros e Avaliadores	Trabalhadores de Traçagem e Montagem de Estruturas Metálicas e de Compósitos	Operadores de Telemarketing
Atletas Profissionais	Farmacêuticos	Forneiros Metalúrgicos (Segunda Fusão e Reaquecimento)	Representantes Comerciais Autônomos
Bombeiros e Salva-Vidas	Químicos	Trabalhadores de Trefilação e Estiramento de Metais Puros e Ligas Metálicas	Psicólogos e Psicanalistas
Vigilantes e Guardas de Segurança	Professores na Área de Formação Pedagógica do Ensino Superior	Operadores de Fornos de Primeira Fusão e Aciaria	Profissionais da Escrita

Source: Elaborated by the author.

APPENDIX B – ADDITIONAL FIGURES AND TABLES OF CHAPTER 3

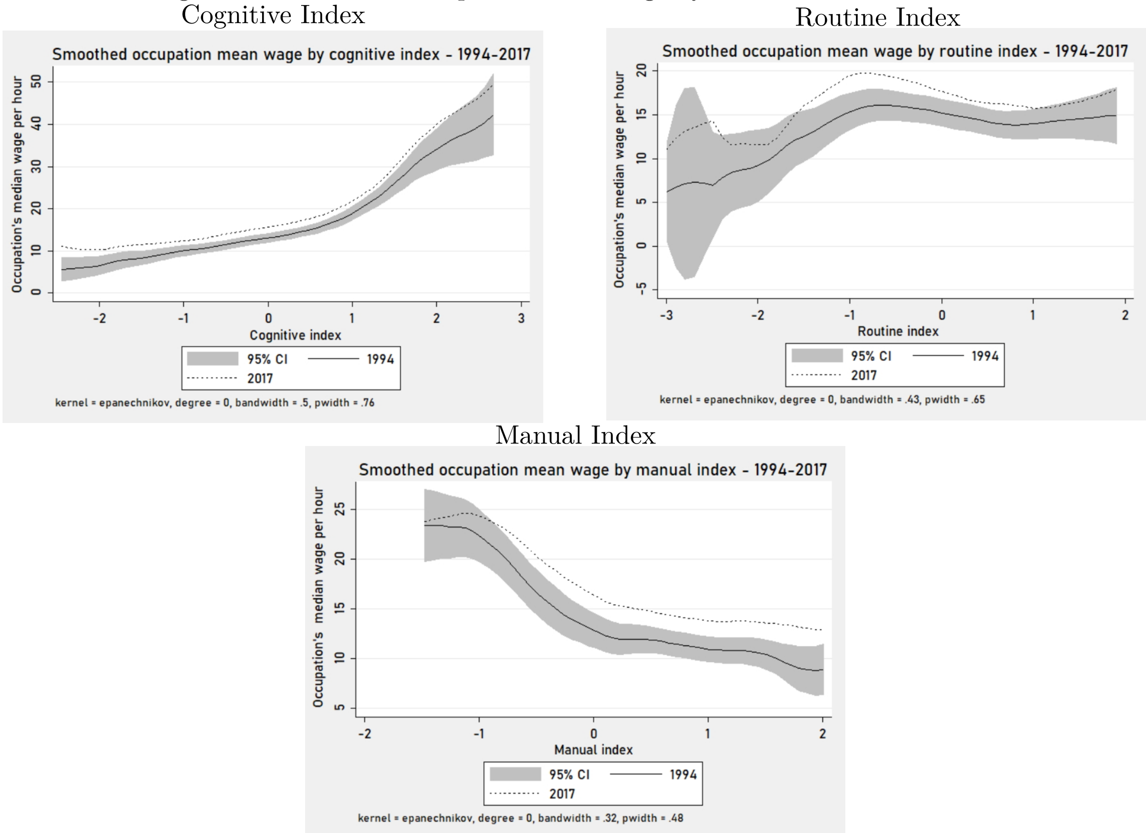
B.1 ADDITIONAL FIGURES OF SECTION 3.4

Figure 9 – Smoothed occupation mean wage by task indexes - 1994-2017

Cognitive Index

Routine Index

Manual Index



Source: Elaborated by the author.

B.2 TABLES OF ACITIVITIES GROUPS

For 2002-2003-2004 panel data, we use section aggregation of activities sectors of CNAE 1995, provided by IBGE, according to the table below. For 2002-2003-2004 panel data, in its turn, we use section aggregation of activities sectors of CNAE 2.0, as showed by Table 21.

Table 20 – CNAE Sections - 1995

Seção	Denominação
A	Agricultura, Pecuária, Silvicultura e Exploração Florestal
B	Pesca
C	Indústrias Extrativas
D	Indústrias de Transformação
E	Produção e Distribuição de Eletricidade, Gás e Água
F	Construção
G	Comércio, Reparação de Veículos Automotores, Objetos Pessoais e Domésticos
H	Alojamento e Alimentação
I	Transporte, Armazenagem e Comunicações
J	Intermediação Financeira, Seguros, Previdência Complementar e Serviços Relacionados
K	Atividades Imobiliárias, Aluguéis e Serviços Prestados às Empresas
L	Administração Pública, Defesa e Seguridade Social
M	Educação
N	Saúde e Serviços Sociais
O	Outros Serviços Coletivos, Sociais e Pessoais
P	Serviços Domésticos
Q	Organismos Internacionais e Outras Instituições Extraterritoriais

Source: IBGE.

Table 21 – CNAE Sections - 2.0

Seção	Denominação
A	Agricultura, pecuária, produção florestal, pesca e aquicultura
B	Indústrias extrativas
C	Indústrias de transformação
D	Eletricidade e gás
E	Água, esgoto, atividades de gestão de resíduos e descontaminação
F	Construção
G	Comércio; reparação de veículos automotores e motocicletas
H	Transporte, armazenagem e correio
I	Alojamento e alimentação
J	Informação e comunicação
K	Atividades financeiras, de seguros e serviços relacionados
L	Atividades imobiliárias
M	Atividades profissionais, científicas e técnicas
N	Atividades administrativas e serviços complementares
O	Administração pública, defesa e seguridade social
P	Educação
Q	Saúde humana e serviços sociais
R	Artes, cultura, esporte e recreação
S	Outras atividades de serviços
T	Serviços domésticos
U	Organismos internacionais e outras instituições extraterritoriais

Source: IBGE.

B.3 TABLES OF FULL RESULTS OF SECTION 3.7

Table 22 – Estimated task price - Full with covariates - 2002-2003-2004 PANEL

	2002-2003-2004 PANEL							
Models	I		II		III		IV	
Cognitive	0.020	(0.0002)***	0.020	(0.0002)**	0.010	(0.0001)***	0.010	(0.0001)***
Routine	-0.003	(0.0001)***	-0.002	(0.0001)***	-0.004	(0.0001)***	-0.003	(0.0001)***
Rout_sq			-0.003	(0.0001)***			-0.002	(0.0001)***
Manual	-0.002	(0.0002)***	-0.003	(0.0002)***	-0.002	(0.0001)***	-0.003	(0.0001)***
Age					0.002	(0.0006)***	0.002	(0.0006)***
Age_quartic					0.000	(0.0000)***	0.000	(0.0000)***
Educ1					-0.011	(0.0012)***	-0.011	(0.0012)***
Educ2					-0.027	(0.0010)***	-0.027	(0.0010)***
Educ3					-0.022	(0.0009)***	-0.022	(0.0009)***
Educ4					-0.025	(0.0009)***	-0.025	(0.0009)***
Educ5					-0.029	(0.0008)***	-0.029	(0.0008)***
Educ6					-0.022	(0.0009)***	-0.022	(0.0009)***
Educ7					-0.021	(0.0007)***	-0.020	(0.0007)***
Educ8					-0.007	(0.0008)***	-0.007	(0.0008)***
Tenure					0.000	(0.0000)***	0.000	(0.0000)***
Dummy2002					-0.107	(0.0001)***	-0.107	(0.0001)***
Dummy2003					-0.073	(0.0001)***	-0.073	(0.0001)***
Metrop. Area					0.002	(0.0011)***	0.002	(0.0011)***
Firm_size1					-0.141	(0.0012)***	-0.141	(0.0012)***
Firm_size2					-0.126	(0.0011)***	-0.126	(0.0011)***
Firm_size3					-0.109	(0.0010)***	-0.109	(0.0010)***
Firm_size4					-0.088	(0.0009)***	-0.088	(0.0009)***
Firm_size5					-0.068	(0.0008)***	-0.068	(0.0008)***
Firm_size6					-0.046	(0.0006)***	-0.046	(0.0006)***
Firm_size7					-0.025	(0.0005)***	-0.025	(0.0005)***
Activ.1					-0.047	(0.0095)***	-0.047	(0.0095)***
Activ.2					0.019	(0.0030)***	0.019	(0.0030)***
Activ.3					0.002	(0.0018)*	0.002	(0.0018)*
Activ.4					0.031	(0.0048)***	0.031	(0.0048)***
Activ.5					-0.011	(0.0021)***	-0.011	(0.0021)***
Activ.6					-0.024	(0.0019)***	-0.024	(0.0019)***
Activ.7					-0.036	(0.0029)***	-0.036	(0.0029)***
Activ.8					-0.004	(0.0024)**	-0.004	(0.0024)**
Activ.9					0.063	(0.0040)***	0.063	(0.0040)***
Activ.10					-0.017	(0.0019)***	-0.017	(0.0019)***
Activ.11					0.006	(0.0041)*	0.006	(0.0041)*
Activ.12					-0.024	(0.0047)***	-0.024	(0.0047)***
Activ.13					-0.009	(0.0035)***	-0.009	(0.0035)***
Activ.14					-0.024	(0.0026)***	-0.024	(0.0026)***
Activ.15					-0.016	(0.0104)*	-0.016	(0.0104)*
Activ.16					0.029	(0.0198)*	0.029	(0.0197)*
Constant					4.951	(0.0201)***	4.952	(0.0201)***
Obs.	17,991,567		17,991,567		17,991,567		17,991,567	
F	4,425.75		3,367.90		2,368.08		2,046.01	
Prob>F	0.00000		0.00000		0.00000		0.00000	

Source: Elaborated by the author.
Note: Standard errors in parentheses. (*) p<0.10, (**) p<0.05 (***). The results are taken concerning the bases of varaibles: year(2004), education (master and Ph.D.), firm size (highest level: more than 1000 employees), activity (agriculture, livestock, forest production, fishing and aquaculture).

Table 23 – Estimated Task Price - Full with covariates - 2012-2013-2014 PANEL

2012-2013-2014 PANEL								
Models	**I**		**II**		**III**		**IV**	
Cognitive	0.053	(0.0002)***	0.051	(0.0002)**	0.032	(0.0002)***	0.031	(0.0002)***
Routine	-0.011	(0.0001)***	-0.005	(0.0001)***	-0.003	(0.0002)***	0.000	(0.0002)*
Rout_sq			-0.009	(0.0001)***			-0.006	(0.0001)***
Manual	0.002	(0.0002)***	-0.002	(0.0002)***	0.004	(0.0002)***	0.002	(0.0002)***
Age					0.018	(0.0002)***	0.018	(0.0006)***
Age_quartic					-0.003	(0.0000)***	0.000	(0.0000)***
Educ1					-0.021	(0.0015)***	-0.017	(0.0012)***
Educ2					-0.018	(0.0009)***	-0.018	(0.0010)***
Educ3					-0.018	(0.0008)***	-0.018	(0.0009)***
Educ4					-0.019	(0.0007)***	-0.019	(0.0009)***
Educ5					-0.025	(0.0006)***	-0.024	(0.0008)***
Educ6					-0.024	(0.0007)***	-0.024	(0.0009)***
Educ7					-0.021	(0.0005)***	-0.021	(0.0007)***
Educ8					-0.020	(0.0007)***	-0.020	(0.0008)***
Tenure					0.000	(0.0000)***	0.000	(0.0000)***
Dummy2012					-0.109	(0.0004)***	-0.109	(0.0001)***
Dummy2013					-0.046	(0.0002)***	-0.046	(0.0001)***
Metrop. Area					-0.005	(0.0007)***	-0.005	(0.0011)***
Firm_size1					-0.134	(0.0008)***	-0.134	(0.0012)***
Firm_size2					-0.119	(0.0007)***	-0.119	(0.0011)***
Firm_size3					-0.100	(0.0007)***	-0.100	(0.0010)***
Firm_size4					-0.079	(0.0006)***	-0.007	(0.0009)***
Firm_size5					-0.056	(0.0005)***	-0.056	(0.0008)***
Firm_size6					-0.034	(0.0004)***	-0.034	(0.0006)***
Firm_size7					-0.015	(0.0003)***	-0.015	(0.0005)***
Activ.1					0.047	(0.0036)***	0.047	(0.0095)***
Activ.2					-0.010	(0.0023)***	-0.009	(0.0030)***
Activ.3					0.011	(0.0029)***	0.011	(0.0018)***
Activ.4					-0.008	(0.0024)***	-0.007	(0.0048)***
Activ.5					-0.023	(0.0023)***	-0.023	(0.0021)***
Activ.6					-0.017	(0.0024)***	-0.018	(0.0019)***
Activ.7					-0.039	(0.0031)***	-0.039	(0.0029)***
Activ.8					-0.016	(0.0028)***	-0.016	(0.0024)***
Activ.9					0.030	(0.0039)***	0.030	(0.0040)***
Activ.10					-0.003	(0.0036)*	-0.003	(0.0019)*
Activ.11					-0.023	(0.0026)***	-0.023	(0.0041)***
Activ.12					-0.015	(0.0024)***	-0.015	(0.0047)***
Activ.13					-0.027	(0.0025)***	-0.027	(0.0035)***
Activ.14					-0.026	(0.0083)***	-0.026	(0.0026)***
Activ.15					-0.024	(0.0033)***	-0.023	(0.0104)***
Activ.16					0.000	(0.0042)*	0.000	(0.0042)*
Activ.17					-0.014	(0.0027)***	-0.015	(0.0027)***
Activ.18					0.001	(0.0103)*	0.001	(0.0103)*
Activ.19					-0.028	(0.0063)***	-0.028	(0.0063)***
Constant					4.793	(0.0099)***	4.797	(0.0099)***
Obs.	30,938,643		30,938,643		30,938,643		30,938,643	
F	11,283.90		8,794.96		6,158.71		5,321.72	
Prob>F	0.00000		0.00000		0.00000		0.00000	

Source: Elaborated by the author.
Note: Standard errors in parentheses. (*) p<0.10, (**) p<0.05 (***). The results are taken concerning the bases of varaibles: year(2014), education (master and Ph.D.), firm size (highest level: more than 1000 employees), activity (Agriculture, Livestock, Forestry and Forestry).