UNIVERSITY OF HOHENHEIM

Hohenheim Discussion Papers in Business, Economics and Social Sciences

MAPPING STRATIFICATION: THE INDUSTRY-OCCUPATION SPACE REVEALS THE NETWORK STRUCTURE OF INEQUALITY

Dominik Hartmann
University of Hohenheim/Fraunhofer IMW Leipzig/University of São Paulo

Cristian Jara-Figueroa
MIT Media Lab Cambridge

Mary Kaltenberg
Maastricht University/Brandeis University

Paulo Gala
Fundação Getulio Vargas

Research Area INEPA

06-2019

Discussion Paper 06-2019

MAPPING STRATIFICATION: THE INDUSTRY-OCCUPATION SPACE REVEALS THE NETWORK STRUCTURE OF INEQUALITY

Dominik Hartmann, Cristian Figueroa, Mary Kaltenberg, Paulo Gala

Research Area "INEPA – Inequality and Economic Policy Analysis"

ISSN 2364-2084

Mapping Stratification: the industry-occupation space reveals the network structure of inequality

Version June 4, 2019

Dominik Hartmann
Department of Innovation Economics, University of Hohenheim, DE
Fraunhofer Center for International Management and Knowledge Economy, DE
Department of Production Engineering, University of Sao Paulo, BR

Cristian Jara-Figueroa
The Collective Learning Group, The MIT Media Lab, Cambridge, US

Mary Kaltenberg
UNU-MERIT, Maastricht University, Maastricht, NL
Brandeis University, Waltham, MA, US

Paulo Gala
Sao Paulo School of Economics, Fundação Getulio Vargas, BR

Abstract: Social stratification is determined not only by income, education, race, and gender, but also by an individual's job characteristics and their position in the industrial structure. Utilizing a dataset of 76.6 million Brazilian workers and methods from network science, we map the Brazilian Industry-Occupation Space (BIOS). The BIOS measures the extent to which 600 occupations co-appear in 585 industries, resulting in a complex network that shows how industrial-occupational communities provide important information on the network segmentation of society. Gender, race, education, and income are concentrated unevenly across the core-periphery structure of the BIOS. Moreover, we identify 28 industrial occupational communities from the BIOS network structure and report their contribution to total income inequality in Brazil. Finally, we quantify the relative poverty within these communities. In sum, the BIOS reveals how the coupling of industries and occupations contributes to mapping social stratification.

Keywords: labor markets, social structure, stratification, economic sociology, wages, inequality

JEL: J31, L0, Z13

1 Introduction

“What do you do for a living?” is probably one of the most common questions asked to strangers at social gatherings. This question partly satisfies the search for common interests and the need for societal placement. Knowing someone’s occupation translates into knowing not only their earnings, but, as sociologists have long argued, also reflects prestige and status within society (Weber, 1922; Blau and Duncan, 1967; Kalleberg and Berg, 1987, Mouw and Kalleberg, 2010). Most approaches in economics distinguish different social classes based on differences in income, consumption, skills, or education. In contrast, a considerable number of sociologists have highlighted the role of occupations as an identifier of social class (ibid.). Less emphasis has been given to the industries that bring people from different occupations together, and to the extent to which different occupations cluster together in these industries. Overlooking work place characteristics is an important shortcoming, because people spend a large share of their time with colleagues in the workspace, sharing knowledge and common interests, and comparing themselves with their work-related peers across occupations within industries. Apart from the social spaces that the workplace impose, different types of industries are associated with different levels of job stability, wages, and social prestige (Campbell, 1960; Tatro and Garbin, 1973; Thaler, 1989; Binder et al., 2016, Brummund and Connolly, 2018). For example, working as an administrative assistant in the federal government is likely to associated with a higher job stability and social recognition than working for a small company in a service industry. In consequence, not only occupations matter to define micro-classes (Grusky and Sørensen, 1998; Weeden and Grusky, 2005), but also the industries in which people work. While there are large wage differentials across occupations, there are also large inter-industry wage-differentials (Krueger and Summers, 1988; Taylor, 1989) and differences in social prestige (Campbell, 1960). The interaction between industries and occupation is part of what determines the wage, social network and social status of a person and structures society into different hierarchically and horizontally clustered subgroups.

In this article, we show that social stratification in modern complex economies is not only about income, education, race, or gender, but also about the type of industries in which different professions work. The coupling of industries and occupation creates different social subgroups with shared interests and knowledge, and thus segments society. The industry in which one works matters as there are industry specific knowledge interests and physical spaces that condition the likelihood of an individuals’ social interaction and common interests. These differences in social interactions and interests are likely to cluster society into a network of different work-related societal groups. We focus our analysis on Brazil because it is known to be a segmented, unequal, and structurally heterogeneous country (Furtado, 1959, 2009) and provides a fine-grained dataset on the occupations and industries in which people work.

Research in sociology has long highlighted the role of social prestige, networks, and occupations for social stratification and inequality (Weber, 1922; Bourdieu, 1986; Weeden and Grusky, 2005 Lin, 2017, Zhou and Wodtke, 2018). Yet most of the sociology research that considers social networks and occupations has important shortcomings. Research on social classes and work structures has tended to focus on the national level (Kalleberg and Berg, 1987) or distinguished only between broad social classes and occupational groups (Marx, 1867; Weber, 1922; Bourdieu, 1984; Erikson et al., 1979). Only recently has research highlighted micro-classes of different disaggregated occupations (Weeden and Grusky, 2005; Weeden and Grusky, 2012). These approaches provide important insights on the social class structure, yet still fail to capture how the interaction of occupations and industries segments and stratifies modern complex economies and societies. Arguably, the network position of people in the productive network, both in term of the occupation and the industry in which people work, strongly conditions their social status, networks, and ability to be active agents of development (Hartmann, 2014).

The relative network position and segmentation of different socioeconomic groups in the economy can also explain part of what is colloquially referred to as the, "Keeping up with the Joneses" effect, which is the tendency of people evaluate their income position relative to their neighbor, and thus, misevaluate their own income position in the greater national society. Research on income inequality has affirmed this effect and shown that the distance between socioeconomic groups is often misperceived. Studies in the United States revealed that people tend to underestimate the severity of income inequality within their country (Norton and Ariely, 2011, Clark and Senik, 2010). Since individuals tend to compare themselves relative to their neighbors, they can increase their happiness from relative income comparisons and not only from consumption (Luttmer, 2005; Guven and Sørensen). Arguably, "Keeping up with the Joneses" effects are also present in the work place. As an example, an engineer earning a relatively high income may feel relatively poor, if her department manager is buying a new car that she cannot afford.

Moreover, while there is a general understanding that society is clustered into a complex network of different economic activities, methods to capture the network structure of industrial-occupational neighborhoods are just starting to be more widely used (Jara-Figueroa et al, 2018). In this paper, we contribute to filling this gap by analyzing a detailed dataset of 76.6 million workers in Brazil to reveal the network structure of inequality and social stratification in the Brazilian Industry-Occupation Space (BIOS). We build a network called the Industry-Occupation space to reveal the relatedness between occupations in terms of the industries in which they frequently co-appear. We then analyze how income, race, gender, and education are distributed across the BIOS, distinguish different work-related socioeconomic groups, and identify local rich and poor within these groups. The remainder of the paper is structured as follows. Section 2 provides a brief literature review on how different types of industries and occupations are connected to different social classes in an economy. Section 3 introduces the data and methods used in this paper. Section 4 presents the Brazilian industry-occupation space and discusses different socioeconomic groups identified by network community detection algorithms. Section 5 provides concluding remarks.

2 Literature on occupations, industries and social stratification

A wide variety of factors condition social stratification. Traditionally, social stratification has been characterized by differences in income, education, consumption and lifestyles, geography, race, and gender. Here, we focus on stratification imposed by the occupations and industries in which people work.

Occupations represent a multitude of social markers: income brackets, human capital, work environment, and skills. The interaction of occupations and industries define specific labor markets that constitute the productive structure, or what some sociologists call, the work structure (Kalleberg and Berg, 1987). Economists, however, typically distinguish social class by their level of income and thus, consumption power. In contrast, sociologists aim to understand the formation of socioeconomic classes in terms of the social conditioning, and the institutionalization of social conditions. Weeden and Grusky (2005) argue that these processes of social conditioning and institutionalization occur at the occupational level. Consequently, they create ‘class maps’ by utilizing detailed occupational data of grouped occupations based on lifestyles, sentiments, demographic composition, and life changes. The divergence of the research focus between economists and sociologists is also notable within the social mobility literature, where economists tend to consider income changes as the main marker of social mobility while sociologists often study social mobility in terms of occupational shifts (Blau and Duncan, 1967; Rytina, 1992; Heckman and Mosso, 2014). Both economists and sociologists, however, tend to agree that occupations should be categorized based on their skills, tasks, and know-how (Kim and Sakamoto, 2008).

Economists often look at the labor market as a singular market where workers match with firms’ labor demands. The ability of a worker to move between jobs, however, does not occur in one market, but rather many labor market pools that are specialized and depend on a variety of characteristics, such as geography, occupation and industry-specific skills, shared institutions, certificates, or educational achievements. In some cases, occupations within an industry are closely related to each other. In other cases, moving between industries, such as from the textile to service industry, may be easier than moving between occupations within an industry.

Clustered networks of related occupations and industries also determine wages and hierarchically cluster subgroups. For instance, industries and occupations related to high revenues, e.g. oil industries or recently digital technologies, tend to have a larger “cake” that can potentially share with its workers, than competitive industries with relatively low profit shares, like e.g. low-tech service industries. However, wages are also dependent on the labor market concentration and buyer-supplier relationships. In this regard, recent work on labor market monopsony and large-scale buyers showed that labor market concentration can reduce wages within certain type of occupations and industries (Azar et al., 2017, 2018; Wilmers, 2018; Rinz, 2018). This implies also that wages are partially determined by the industry and related labor market within that industry.

Moreover, the relatedness between different types of occupations and industries can arguably be a good indirect measure of social interactions or separation between different socioeconomic groups. The workplace provides opportunities for frequent interactions with other people coming from similar knowledge backgrounds. Frequent interactions between people build trust and empathy, and thus facilitate in-depth knowledge transfer (Coleman, 2000). There may be weak links and knowledge exchange between people from very different occupations and industries, but more in-depth knowledge transfer, shared interests, and imitative behavior tend to happen between people from similar occupations and industries. In consequence, the clustering or distance between different types of occupations and industries can arguably reveal the complex network structure of segmentation and stratification within an economy and society.

It must be noted that different types of industries are also associated with different types of social prestige, wages, and job stability (Campbell, 1960; Tatro and Garbin, 1973; Binder et al., 2016, Brummund and Connolly, 2018). The prestige of industries can differ across regions and over time. For instance, jobs in information or digital technology have significantly increased and become the elite work force in the US (Binder et al., 2016). In contrast, a stable and relatively well-paid job in the public administration continues to be desired by the middle to upper class in Brazil, and are fiercely fought for in public exams. Again, this points to the close linkage between occupation and industries as twin forces for social stratification and inequality.

Research in development economics and institutional economics has highlighted the role of industries in the evolution of inequality across and within economies (e.g. Furtado, 1959; Engerman and Sokoloff, 1997; Hartmann et al., 2017, 2019a, 2019b). In the context of former colonial economies, the socioeconomic stratification and inequality between different occupational and industrial groups tend to be exacerbated due to a variety of historical and socioeconomic reasons. The type of extractive institutions established in a colony affects the future level of inequality, both in terms of income levels between the colonizer and local economy, and the social relations among the local population (Furtado, 1959; Acemoglu and Robinson, 2012). In Brazil, for instance, the establishment of exploitative colonial production schemes between the 16th to 19th centuries led to a high dependence on commodities, very high levels of income inequality, and a strong differentiation of social classes in Brazil (Freyre, 1933; Furtado, 1959). Moreover, Latin American economists have highlighted the unequal diffusion of technology and productivity across different productive sectors and related occupations in the formation of Latin American economies (Prebisch, 1949; Furtado, 1959). Economists have argued that some industries benefit from significant technological advances and levels of productivity, such as modernized large-scale agriculture or some manufacturing industries, while a sizeable number or low-productivity sectors, such as street commerce or subsistence agriculture, have produced a continuous surplus of labor, keeping wages low and inequality high (Rodrik, D. and McMillan, M. 2011, Rodrik 2016).

Social stratification, economic structure and inequality in Brazil

With a Gini of 51.3 according to the World Bank development indicators in 2015, Brazil is among the most unequal countries in the world, with historically large gaps in income and social status between different occupations and industries. In this regard, Gilberto Freyre (1933) highlighted the social distinction between "masters and slaves" that developed during the formation of Brazil. More recent work in institutional economics argued that economic specialization in exploitative industries, such as large-scale sugar or coffee plantations or mining activities, has been associated with exploitative institutions and perpetuated high levels of income inequality that continued to this day (Engerman and Sokoloff, 1997).

Despite a considerable diversification into some manufacturing and some knowledge-based service industries, such as aerospace, cars, petrochemical, electronic, and finance industries, large segments of the Brazilian economy and society are working in relatively simple and exploitative economic activities such as agriculture, mining, or simple services (Gala et al. 2018; Hartmann et al., 2016, 2017, 2019). The high level of structural heterogeneity between high and low productivity economic activities continues to create high levels of inequality and worlds apart between different socioeconomic groups. Often, knowledge-intensive jobs like bank managers or aerospace engineers regionally coexist with jobs like cleaning and housing services and different types of blue-collar jobs in construction, agriculture, and various manufacturing industries. But those jobs are socially far apart from one another, not only in terms of income, but also in terms of the industry related knowledge, frequency of social interaction, and socioeconomic milieus.

It is noteworthy that Brazil has a strong tendency towards conspicuous consumption. Besides the standard class evaluation of based income strata (e.g. by the Brazilian Institute of Geography and Statistics), another renowned classification is the "Criterion of Economic Classification" of Mazzon and Kamakura (2016) that puts emphasis on consumption and living standards. This classification captures the living standard, based on taking, for example, the number of housemaids, bathrooms, cars or televisions of the household into accounts. Less emphasis has been put on classifications based on differences across occupational groups. Several studies on income inequality in Brazil have highlighted differences in the productivity of different types of industries, the effect of differences in education and the association between race discrimination, slavery and inequality (Freyre, 1933; Bourguignon et al., 2007; Tavares and Menezes-Filho, 2011; Rodriguez-Castelán et al., 2016; Fujiwara et al., 2017), but relatively few works have analyzed the stratification associated with occupations on a disaggregated level (e.g. Maia and Sakamoto, 2016; Bartoncelo, 2016). In some cases, strikes or demonstrations of single occupational groups, such as the "Movimento Sem Terra" of landless, agricultural workers, have drawn attention to particular occupations. Moreover, the enormous class divide between activities associated to the poorer classes, such as housemaids, streets sellers, and construction workers, and occupations associated to the rich and powerful, such politicians, judges, managers, or media stars is omnipresent in the daily life. Yet a more comprehensive empirical picture of the socioeconomic stratification of the Brazilian economy and society which distinguishes different industrial-occupational groups is largely missing. A mere top-down classification from rich to poor, based on different income layers or consumption schemes may omit the complex social differentiation and stratification between different types of industries and occupations.

3 Data and Methods

Recent interdisciplinary methods on bipartite networks have helped to analyze relatedness between different knowledge fields (Hidalgo et al. 2018)—such as scientific disciplines (Guevara et al., 2016, 2017), patents (Alstott et al., 2017) or industries (Hidalgo et al., 2007, Hartmann et al., 2017, Jara-Figueroa et al, 2018). Here we make use of these methods to understand how occupations are linked with each other via the industries they share. Moreover, we explore how different socioeconomic characteristics such as income, gender, race, or education are distributed across the resulting Industry-Occupation Space.

We use Brazil's Annual Relation of Social Information (RAIS) (Cardoso, 2007) for 2006 to 2013 to create an industry-occupation space. RAIS is an administrative register compiled by the Ministry of Labor (MTE) based on information offered compulsorily by all formally registered, public or private companies in the country. MTE estimates that RAIS is annually declared by 98% to 99% of officially existing firms (Cardoso, 2007). The variables in RAIS are available at the municipality level, which makes it the most important source of information on the formal labor market dynamics in the country. The variables in RAIS are collected annually and include demographic, occupational, and income characteristics of employees, as well as labor force movement (hiring and firing). RAIS includes fine-grained information about individual workers in Brazil, including 5,560 municipalities, 2,500 occupations, and 585 industries for more than 30 million workers each year. Occupations are classified according to the Brazilian Occupation Classification (CBO). Here, we use 4-digit level, which gives us a total of 600 different occupation categories. This fine-grained dataset allows us to reveal the social clustering and stratification in the Brazilian industry-occupation space.

The industry-occupation space is a weighted bipartite network connecting industries with their required occupations. This network reveals socioeconomic relatedness between different occupational groups. We formalize the industry-occupation space as follows. Let X_{io} be the number of people working in occupation *o* in industry *i*. We say that *i* is connected to *o* when there are more employees performing *o* in *i* than expected merely by the size of *o*. To formalize this notion, we use the revealed comparative advantage RCA_{io} of occupation *o* for industry *i*:

$$RCA_{io} = \frac{X_{io}}{\sum_{o'} X_{io'}} \Bigg/ \frac{\sum_{i'} X_{i'o}}{\sum_{i'o'} X_{i'o'}}$$

When the share of *o* in *i* is greater than of *o* in the entire population (i.e. when *RCA > 1*) then that occupation *o* is a relevant occupation for industry *i*. This procedure handles the fact that there may be occupations that are present in an industry only because the industry is large, and not because they are a key occupation in this industry. By using RCA, we make sure that the connections between an industry and an occupation indicate the role this occupation plays in this industry.

Let M_{io} be the adjacency matrix of the network that connects occupations and industries. M_{io} is such that:

$$M_{io} = \begin{cases} 0, & RCA_{io} < 1 \\ 1, & RCA_{io} \geq 1 \end{cases}$$

M_{io} is the adjacency matrix of a bipartite unweighted undirected network.

This bipartite network represents the distance (or relatedness) between occupation *o* and occupation *o'*. Let $n_{oo'} = \sum_i M_{io} M_{io'}$ be the number of industries that hire both *o* and *o'* (with *RCA > 1*), and $n_o = \sum_i M_{io}$ the number of industries that hire *o*. The *distance* between *o* and *o'* is defined as:

$$\phi_{oo'} = \min\left\{\frac{n_{oo'}}{n_o}, \frac{n_{oo'}}{n_{o'}}\right\}$$

The distance index can be interpreted as an exposure index that depends only on industry structure. In other words, it captures the odds that a random person from o and one from o' work in the same industry, after taken into consideration the effects of the size of the industries and the frequency of occupations in each industry. This "structural" measure enables us to discuss the part of stratification that is due to occupations appearing in the same industry.

The matrix $\phi_{oo'}$ represents an undirected weighted network. Two occupations in this network are connected when there is a significant number of industries that hire them both. For example, in the case of Brazil, the occupation "mathematics professor" is connected with the occupation "economics professor" because more than 82% of the industries that hire mathematics professors also hire economics professors. In this particular case, the weight of the link is 0.82. Thus, the weights of the links show us to which extent occupations tend to share similar industries and give us a proxy network about how the Brazilian labor market is structured into different socioeconomic spaces.

Clusters of related occupations

We use network community detection algorithms to identify clusters of related occupations. Highly connected occupations based on the industrial structure provide an understanding of how occupations are linked via labor inputs reflecting which occupations are likely to work alongside each other. We take two approaches to identify network communities: the Peixoto stochastic blockchain detection model and the Louvain method. It must be noted that community detection models are imperfect and robust detection with clear benchmarks are still being developed (Fortunato et. al., 2016). Despite this limitation within the current state of network science, these models are still useful for identifying core-periphery structures. We use two different types of detection models to identify clusters that are within the densely connected core (stochastic block method, (Peixoto, 2014)) and the sparser connected periphery (Louvain detection method, (Blondel et. al, 2008)). The stochastic block community detection model's strength lies in its ability to identify core-periphery structures in large datasets with high accuracy and varied community degrees across the network (Fortunato et al. 2016). We identify the core periphery structure in the Brazilian Industry-Occupation Space with Peixoto's method which uses Markov Chain Monte Carlo (MCMC) to modify block membership of each node to find the best partition. To further identify the periphery clusters with higher resolution, we use the popular Louvain community detection method, which searches for modularity partitions and evaluates the gains of each node that is added or removed in a community and continues this process until the local maxima of modularity is achieved. In sum, community detection is a valuable resource to analyze large complex networks by synthesizing highly connected nodes into analytical groups. In this article it allows us to investigate groups of occupations that are connected to each other due to their shared industries.

Contributions of the occupational cluster to inequality in the labor market

To separate the contribution of the occupational clusters in the industry-occupation space to the overall level of income inequality in Brazil, we use the decomposition of the Theil index for different communities (Novotný, 2007; Bourguignon et al., 2007; Xie et al., 2016). The decomposition separates the inequality arising from wage differences within each group of occupations and the deviations of the average wage of each community with respect to the national average wage. These two quantities can be interpreted as *within community inequality* and *between community inequality*, respectively.

Finally, we calculate a set of further socioeconomic characteristics associated with different occupations and occupational groups. In particular, we focus on the number of people the average level of literacy, the Blau diversity of race, the most frequent race and the male / female ratio in the occupations, in the respective occupations and BIOS groups.

4 Results

In this section, we present the Brazilian Industry-Occupation Space (BIOS) and reveal the distribution of race, gender, education, and income across the BIOS.

Core-Periphery Structure of the BIOS

The industry-occupation space captures the differentiation of society into different work-related groups (see Figure 1). We built the Brazilian Industry-Occupation Space (BIOS) using data from 2006 to 2013, considering only occupations and industries that had more than 100 employees. Since the structure of the BIOS is largely stable across a period of several years, we use the data for the entire period between 2006 and 2013. In order to visualize the network, we start with a minimum spanning tree and then populate the network edges with all links that have a weight ($\phi_{oo'}$) higher than 0.55 so that the average degree of the visualized network is between 3 and 4. This allows us to produce a skeleton and visually analytic network (following the same algorithm as in Hidalgo et al, 2007). The results are shown in Figure 1. It must be noted that this filtering process is only used for visualization purposes. To identify network communities, indicated by different node colors, we used the full network.

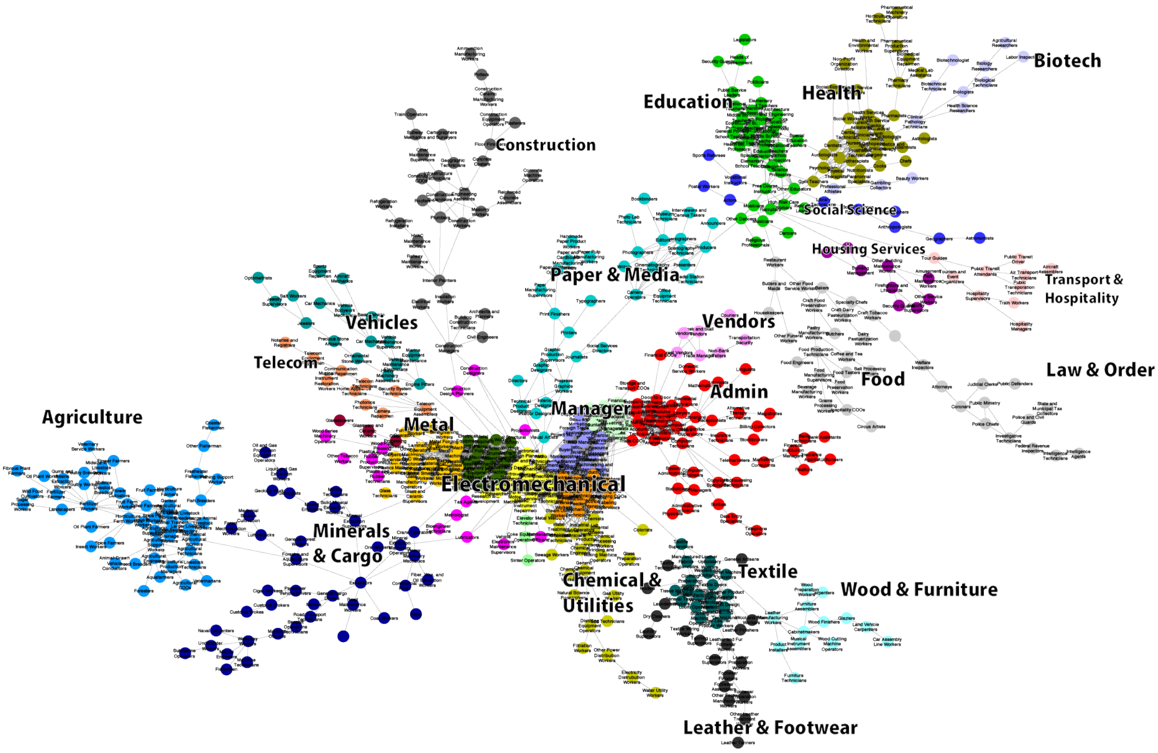

Figure 1. The Brazilian Industry-Occupation Space. Each node presents one of 600 CBO occupations and is colored according to 28 network communities identified by a combination of Peixoto (2014) and Louvain community detection algorithms. Links between the nodes depict to which extent 600 occupations co-appear in 558 different types of industries.

The skeleton network structure reveals the core-periphery structure of the Brazilian industry-occupation space. There is a dense web of occupations in the core that are related to electromechanical, metal, chemical and business management industries. There is strong interconnectedness between firm management/ administrations and the productive activities in the core of the BIOS.

In the periphery of the BIOS, we see several economic activities supporting the core: infrastructure (transportation, maintenance, maritime workers and construction), human development (education and health), services (technicians, media, designers, public justice, food/restaurants, retail, and security), law, and agriculture. Relatively densely connected groups within the periphery are occupations associated with agriculture, textile, health and education industries, respectively. The average path lengths of the occupations in the core of the BIOS are shorter and the network connectivity is higher than in the periphery of the BIOS (See also Figure S1 and S2 in the appendix). This means that occupations at the core, such as business admins or managers, are linked with many other occupations through a varied set of industries that require these types of occupations. Conversely, several occupations in the periphery, such as train operators, book binders, or aircraft assemblers, can only be found within particular specialized industries. Thus, these more central occupations, tend to also have access to a more varied set of knowledge from different industries and occupations, and thus take a more central position in the society. A manager, for example, would be exposed to a diverse set of occupations over the course of their career, while a more specialized occupation such as a medical doctor, would be exposed to a less diverse set of occupations.

Network Communities in the BIOS

We applied stochastic block methods (Peixoto, 2014) and Louvain community (Blondel et al., 2008) detection methods to identify the most connected occupational groups in the BIOS. The Peixoto algorithm is better able to identify groups within the core of the network while the Louvain method can distinguish the periphery clusters with better accuracy. The Peixoto distinguished 12 distinct clusters, while the Louvain revealed 21 clusters. Taken together, we can distinguish 28 occupational clusters, based on the connectivity between different groups of occupations and the distance to other groups of occupations.

It must be noted that these 28 occupational clusters are significantly different from the occupational groups in the Brazilian Occupation Classification (CBO) (Ministério do Trabalho e Emprego, 2010). The occupational classification of Brazil is mainly based on the skills that are necessary to effectively do a job, such as cognitive or manual skills. Yet, in practice, the expected income, social prestige, and skills are also commonly dependent on the particular industry in which one works (Campbell, 1960). Not only the occupation per se, e.g. being an administrative assistance, assembly or service worker, but also the more precise industry in which a person works, condition different socioeconomic groups.

To test the differences between our cluster solution and the CBO occupational groups we calculate Adjusted Rand Indices (ARI) for several different CBO aggregation levels. The ARI is a standard index that quantifies how similar two classifications are, after adjusting by chance, where 0 means that they are as similar as given by chance, and 1 means that they are more similar than what can be expected by chance. We can observe low Adjusted Rand Index (ARI) values of 0.09, 0.19, and 0.10 comparing our 28 occupational clusters in the BIOS solution with occupations 187, 45, and 9 grouping identified on the 1, 2, or 3-digit level of the CBO. In addition, other cluster solutions, using either Peixoto or Louvain algorithms lead to low ARI indices (See Table S1 in the appendix). Thus, the grouping resulting from the network communities in the BIOS is significantly different from the grouping of occupation in the standard Brazilian Occupation Classification. This is the case because our BIOS clusters do not consider only occupation-related skills, but also the type of industries shared by occupations and thus how different types of occupations and industries cluster together. To a certain extent, this captures how shared knowledge, skills, interests and social networks not only span within one occupation, but also across related or shared industries.

Next, we analyze the resulting 28 occupational groups in more detail. Table 1 shows the number of people, the average wage, literacy, diversity of race, percentage of females in each group, as well as the contribution of each group to the Theil income inequality in Brazil. As expected, we find that directors and senior managers, as well as engineers have the highest average income, workers have the lowest, and technicians have intermediate levels of average income. Yet the BIOS also allows for a more fine-grained distinction between different types of technicians and workers, according to the particular type of occupation and industry they are working. The particular industry in which different types of technicians are working, e.g. in materials, utilities, chemical, electromechanical industries, or telecommunication, is associated with different levels of wage, literacy, diversity and the percentage of female, and identifies distinct socioeconomic groups. Thus, industries structure society into different subgroups. For instance, while technicians and professionals in utilities and chemical industries are mainly male, with a medium level of education, and a high wage, technicians and professionals in health and biotechnology tend to be female, earn slightly less income, and have a higher education level than the previous group.

It is noteworthy that while the groups “Directors” and “Engineers” have the highest average wages, technicians in administration and technicians in education, have the highest total contribution to income inequality. The reason for this is that jobs in administration are much more frequent than CEOs or engineers. In other words, while occupations traditionally associated with the upper class, such as managers and engineers, do have the highest wages, occupations associated within middle or middle-upper class contribute the highest amount of inequality to the Brazilian labor market. Occupations with various general workers add relatively little to overall income inequality in Brazil simply because they are the vast majority of the working population.

Subsequently, we analyze the distribution of gender, race, education, and income across the BIOS in more detail.

Distribution of women in the BIOS.

Figure 2-A shows a marked distribution of the percentage of female in the BIOS. Health, education, administration, and occupations in the garment and fashion industries have a significantly higher percentage of female workers than male-dominated occupations in extractive, agricultural, and manufacturing industries. For instance, 88-95% of people that are working in occupations such nutritionists, social workers, housekeepers, psychologists, and dental technicians are women (see Table S2 in the appendix). In contrast, 99% of people working as construction workers, excavators, general cargo drivers, and fishermen are men. Thus, the BIOS clearly captures the segregation of the working world into industries and occupations that are dominated / preferred by either men, women or both genders. As expected, occupations that tend to be associated with manual and technical skills are dominated by men, and activities associated to social, interactive and cognitive skills are dominated or at least have a larger percentage of women. A promising line of follow-up research would analyze the overlay skills.

Distribution of education in the BIOS.

Next, we analyze the distribution of education—measured by the average level of literacy—in the BIOS (see Figure 2-B). Occupational groups in education, health, management and administration, law, and electromechanical industries tend to have very high levels of education. Several occupation groups in the periphery of the BIOS, such as occupations in agriculture, construction, wood, and textiles tend to have low levels of education.

Table 1. Characteristics of the BIOS network communities in 2013. The communities were labeled manually according to the occupations present in them.

Net group	Description	n_people	Av. Wage	Av. Literacy	Blau Index	% Women	Theil Within	Theil Between	Total wage share	Contribution to total Theil inequality
24	**Directors - Administration, Organization & Planning**	10915315	3507	7.75	0.45	0.46	0.53	0.67	0.07	0.08
20	**Engineers - Electromechanical**	3057213	3306	7.24	0.50	0.09	0.48	0.62	0.02	0.02
23	**Technicians - Utilities & Chemical**	4050396	2579	6.45	0.56	0.13	0.44	0.37	0.02	0.01
19	**Teachers, Artists, Politicians - Education**	38338024	2231	7.98	0.54	0.64	0.34	0.22	0.14	0.08
27	**Department and operational managers**	11889974	2139	6.77	0.54	0.26	0.53	0.18	0.05	0.03
8	**Researchers - Social Sciences**	1235316	2123	7.51	0.57	0.33	0.25	0.17	0.00	0.00
18	**Health Professionals**	21974921	2049	7.29	0.55	0.72	0.38	0.14	0.08	0.04
17	**Technicians - Administration**	62531536	2048	7.30	0.52	0.59	0.50	0.14	0.21	0.13
4	**Technicians - Media & Paper**	3222690	2040	7.05	0.51	0.28	0.38	0.13	0.01	0.01
11	**Technicians - Materials**	783327	2038	6.40	0.53	0.18	0.34	0.13	0.00	0.00
6	**Biotechnology – Different professions**	1214688	1999	7.07	0.58	0.72	0.69	0.11	0.00	0.00
0	**Services (Food, Hospitality, Law & Order)**	16067781	1999	6.23	0.56	0.45	0.72	0.11	0.06	0.05
14	**Coke & Sinter – Different professions**	211625	1827	6.49	0.57	0.11	0.18	0.02	0.00	0.00
25	**Technicians & Workers - Metal & Mechanical**	7501011	1782	6.24	0.53	0.13	0.22	0.00	0.02	0.00
26	**Technicians & Workers - Electromechanical**	14737683	1779	6.30	0.54	0.22	0.33	0.00	0.05	0.01
16	**Technicians - Telecom**	1523113	1773	6.67	0.57	0.06	0.21	-0.01	0.00	0.00
21	**Workers - Metals & Materials**	3311716	1592	5.92	0.50	0.11	0.17	-0.12	0.01	0.00
13	**Transport & Tourism**	5481346	1569	6.01	0.56	0.16	0.16	-0.13	0.01	0.00
9	**Workers - Natural Resources**	20895860	1519	5.66	0.59	0.08	0.26	-0.16	0.05	0.01
5	**Workers & Technicians - Vehicles & Jewelry**	3193939	1517	5.96	0.55	0.04	0.20	-0.16	0.01	0.00
1	**Construction – Different professions**	18694148	1493	5.27	0.62	0.05	0.35	-0.18	0.05	0.01
3	**Workers - Wood & Furniture**	2708294	1255	5.94	0.54	0.07	0.16	-0.35	0.01	0.00
15	**Workers - Glass**	111962	1207	6.04	0.54	0.17	0.10	-0.39	0.00	0.00
7	**Agriculture – Different professions**	12489903	1187	4.42	0.60	0.13	0.28	-0.41	0.02	0.00
10	**Workers - Vendors**	32343269	1066	6.62	0.54	0.54	0.19	-0.52	0.06	-0.02
12	**Workers - Security / Building Maintenance**	27290282	914	5.32	0.61	0.43	0.10	-0.67	0.04	-0.02
22	**Supervisors - Garments**	5937401	903	5.90	0.50	0.70	0.10	-0.68	0.01	0.00
2	**Workers - Garments**	3391853	877	5.52	0.52	0.53	0.08	-0.71	0.00	0.00

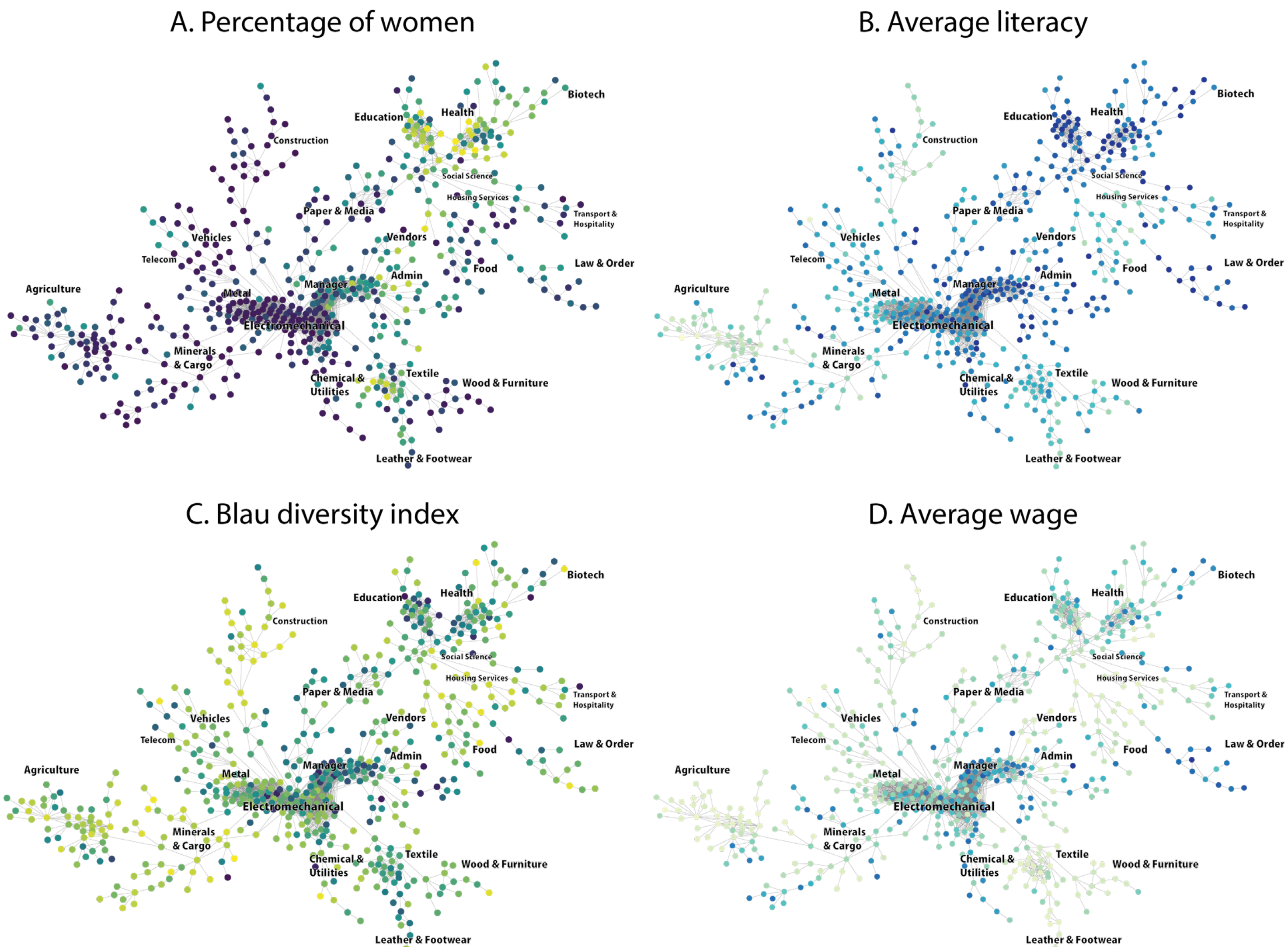

Figure 2: Gender, literacy, race, and wage across the BIOS. **A.** Overlay of the **percentage of women** in the occupations of the BIOS. Blue indicates a higher percentage of men, yellow a higher percentage of women. **B. Average literacy.** Dark shades of blue illustrate higher values, more transparent nodes lower literacy levels. **C. Blau Diversity Index**. Dark blue illustrates a low level of racial diversity. Yellow indicates a high level of racial diversity. Values are calculated with R package Diverse (Guevara et al., 2016). **D. Distribution of wages**. Dark shades of blue illustrate higher values.

Racial diversity in the BIOS.

Next, we analyze the distribution of racial diversity across the BIOS (See Figure 2-C). Occupations associated with a relatively high level of diversity and a high share of black workers tend to have relatively low wages and education levels. In contrast, occupations associated with relatively low levels of diversity and a high percentage of white workers tend to be associated with higher average wages and education levels. For instance, workers in elite occupations, such as financial institution manager, computer engineers, pilots, or health science researcher tend to be white. In contrast, the highest share of workers in occupations such as tobacco, salt, fiber wax, oil, liquid and gas extractions extraction workers are black (see Table S3 in the appendix). Moreover, occupations with a high level of diversity tend to be in the periphery of the BIOS. Interestingly, though racially less diverse occupations are not only slightly clustered in management and administration groups, but are also scattered across the entire BIOS. It seems the history of slavery has persisted in the distribution of racial diversity in the BIOS.

Wage Inequality in the BIOS.

Finally, we analyze the wage distribution across the BIOS in detail. Figure 2-D shows that occupations in the productive core of the BIOS, as well as occupations in health, education, law and order have significantly higher average wages (colored in dark blue) than most occupations in the periphery of the BIOS (light blue).

Interestingly, we can observe some occupations across the BIOS that have significantly higher incomes than their immediate network neighbors and industrial occupational network community. Moreover, there are some occupations that have relatively high wages (e.g. in the administration and management occupations) yet are relatively poor in comparison with some network neighbors that are very rich. This relative poverty within industrial-occupational groups may explain a lack of connection with poor strata outside their own network group. In this regard, it must be noted that many people of higher income strata tend to under-evaluate their own income and focus on their relative deprivation in comparison with a small fraction of very rich individuals. In order to get a realistic overview about wages in Brazil, Figure 3 plots the distribution of wages paid in Brazil to 76.6 million workers in the years 2006 and 2013. We can observe that salaries below 1000 Brazilian Real (=440USD in 2013) are very frequent, salaries between 1000 to 3000 BRL (440-1320USD) are much less frequent, while salaries above 3000 BRL are earned by a limited number of people. Only 5.7% of the salaries are above 5000 BRL.

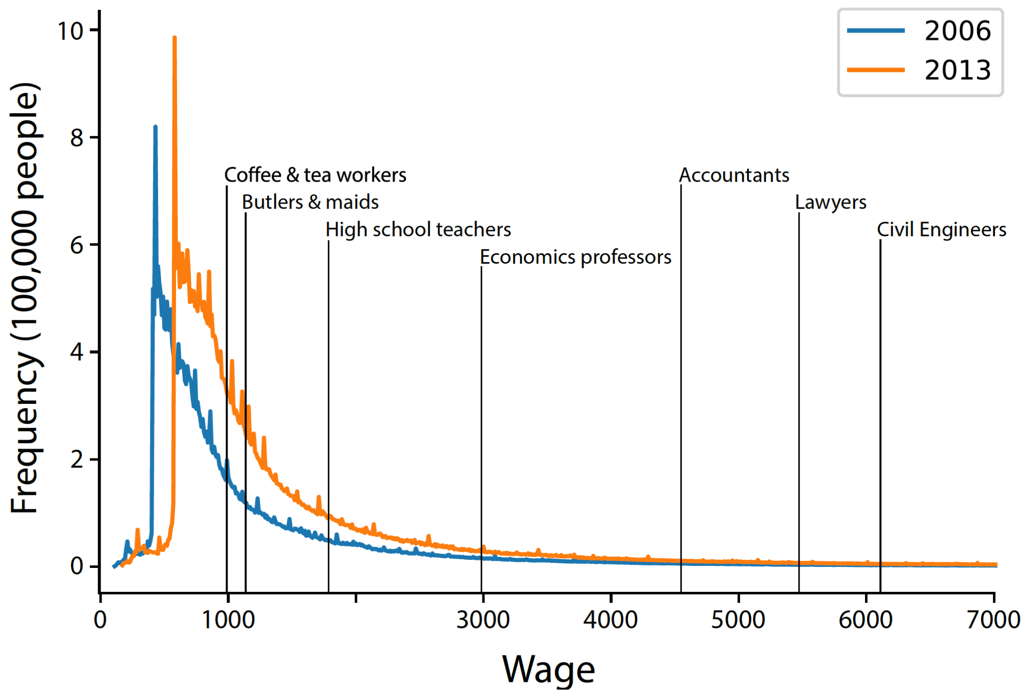

Figure 3. Frequency distribution of wages in the Brazilian formal economy. Black lines indicate the average income of example occupations in 2013.

It is likely that people tend to compare their salaries and living standards rather with people who work in similar occupations and industries—in other words BIOS network neighbors—than with people that work in distant parts of the Industry-Occupation Space. Figure 4 plots the positioning of occupations within their "local occupational network groups" against the positioning in the total income distribution of all occupations. Being relatively poor within the groups is highly correlated (corr=0.815) also with being relatively poor at the national level. Yet there are also differences with respect to the extent to which somebody is relatively above (/below) their own group or the total average wage. On the top left quadrant of Figure 4 are occupations— such as sales supervisors, electromechanical technicians and visual artists— that are above the average national wage, but below the average wage of their industrial occupational cluster. On the bottom right are occupations that are below the total average, but relatively well-off within their network group, such as textile or agricultural supervisors. On the bottom left are occupations that are both below the national wage average as well as their groups average. On the top right are occupations whose workers' average income is both above the national average wage as well as the average wage within their occupational-industrial group.

Figure 4. Average wage of an occupation in comparison to the average wage it is network community (x-axis) and the total average wage in Brazil (y-axis).

Among the occupations with the highest relative advantages within their own group feature high-level public administration occupations, such as ministers, public defenders, and federal revenue inspectors, as well as occupations associated with the mining and deep-sea petroleum industry, such as mining engineers and COO or marine engineers (see Table 2). Among the occupations with the largest disadvantages within their own group feature creative jobs, such as technicians in audio, video and film, administration jobs, such as accounting assistants and secretaries, as well as service jobs such as housekeepers and restaurant workers.

TOP TEN OCCUPATIONS WITH REGARD TO WITHIN GROUP RELATIVE RICHNESS

Rank - within advantage	BIOS community	Occupation	Within community advantage	Av. wage	Advantage to total av. wage
1	0	Public Ministry	239%	22608	249%
2	17	Magistrates	227%	21081	239%
3	0	Federal Revenue Inspectors	195%	14638	205%
4	0	Public Defenders	193%	15890	203%
5	0	Intelligence Agents	190%	13989	200%
6	9	Marine Engineers	184%	12065	167%
7	9	Geologists	178%	11563	161%
8	26	Mining & Quarrying COOs	177%	17917	176%
9	0	Police Chiefs	177%	12450	187%
10	9	Pilots	170%	11314	152%

BOTTOM TEN OCCUPATIONS WITH REGARD TO WITHIN GROUP RELATIVE POVERTY					
Rank - within disadvantage	**BIOS community**	**Occupation name**	**Within community disadvantage**	**Av wage**	**Disadvantage to total av. wage**
595	20	Audio Technicians	-111%	1293	-50%
594	23	Colorists	-107%	1016	-71%
593	20	Video and Film Editors	-104%	1338	-43%
592	24	Accounting Assistants	-104%	1453	-37%
591	17	Secretaries	-103%	804	-91%
590	20	Other Mechanical Fitters	-99%	1515	-39%
589	6	Gambling Collectors	-99%	761	-89%
588	0	Craft Tobacco Workers	-98%	826	-88%
587	17	Telemarketers	-98%	823	-86%
586	6	Beauty Workers	-98%	801	-88%

Table 2. Top / Bottom Ten occupations with respective to their relative within network group wage advantage / disadvantage.

5 Discussion and concluding remarks

Measuring social stratification has often been done in terms of income, education, sentiments, or general skill composition (service, routine, cognitive, etc.). While there are varying perspectives of how to analyze an individual's position within a social structure, we took the perspective of utilizing a person's position based on their occupation, while also including the industry in which they work. As Weeden and Grusky have argued, "occupations shape behavior through additional sociological forces of self-selection, differential recruitment, socialization, and interactional closure, all of which become activated in the context of institutional categories" (p.142, 2005). Our approach follows this perspective as a way to analyze social stratification, yet we defined a narrower social position based on work that not only identifies one's occupation but also includes the industry. This interaction between industry and occupation is important as it is not only the particular job title that an individual holds, but also the unique knowledge of an industry that further constrains their position within the labor market. Industries are a site of stratification, just like occupations and education.

We believe the BIOS provides a map of relative income groups where the complex network structure reveals where there are shared work experiences with occupations and industries. Thus, the industry-occupation space helps reveal who are work-related neighbors. In this regard, a "comparative neighbor" in the industry-occupation space is someone who has a similar job in a similar industry. Arguably, people tend to compare their income with the income of people with similar jobs and industries to a greater extent than to the income of people from very different industries and occupations. The perception of one's income and societal position is influenced by one's work environment, and thus positioning in the industry-occupation space.

We used network community detection algorithms to identify 28 different network communities sharing similar occupations and industries. Moreover, we showed that race, gender, education and income show different distributions across the BIOS. While some parts of the BIOS are dominated by women, others are dominated by men. Moreover, while some parts of the BIOS have high levels of average income and/or high levels of income, other parts of the BIOS have high levels of inequality between some well-paid professions occupied mainly by white men, and low paid occupations occupied by a more racially diverse work force. Finally, we showed that some occupations are relatively well-paid in comparison with the entire workforce, but relatively deprived within their industrial occupational group.

One may ask to which extent the results and industry-occupation space presented in this article is unique to the case of Brazil or depicts typical structures that can also be found in other economies and societies. Several of the findings are not unique to the Brazilian case. Firstly, increasing the division of labor is a widespread phenomenon that is likely to increase the socioeconomic differentiations of different socioeconomic groups. Secondly, we may find a certain disconnection between education and health industries, and the productive core in most countries. This kind of structure can also be found in the United States, which has an interconnected education and health industry, but both of those industries are separated from core activities (Kaltenberg and Hidalgo, 2018). Other countries with stronger vocational systems, such as Germany, may have more overlaps between the education sector and productive industries. Thirdly, a certain disconnection of natural resource driven economic activities, such as agriculture and mining, with the rest of the industry-occupation space can also be found in other countries (Kaltenberg and Hidalgo, 2018).

However, there are also differences. For instance, other economies may have a more diversified and sophisticated service sector, including large business service sectors (Kaltenberg and Hidalgo, 2018). Moreover, the pronounced segmentations of race, education, and income arguably reflects the exploitative historical formation of the Brazilian economy and society (Freyre, 1933; Furtado, 1959; Engerman and Sokoloff, 1997). For instance, the large importance and separation of different types of resource-based activities into separated peripheral parts of the BIOS reflects the large spatial size and natural-resource wealth of Brazil. Hopefully, subsequent comparative analysis of industry-occupation spaces may shed more light on how different historical trajectories and industry specialization affect the social stratification of societies.

There are a couple of additional limitations that need to be mentioned and addressed in subsequent research. Using the RAIS database allows us to analyze the entire formal Brazilian labor market. Yet this also means that this analysis does not include the large informal sector in Brazil. However, we expect that including the informal market may not substantially affect the main findings of this paper. It is likely that the size of the periphery would expand, and some of the richer strata earning even higher relative incomes. The fundamental structure and industrial-occupational groups, though, are not likely to change. However, there is a need for a more sophisticated classification of services. The lack of more fine-grained disaggregation of service industries may explain why we identify one larger service group that is comprised of both simple jobs, like housemaids and butchers, as well as professions in the law and order sector, such as attorneys, public ministry employees, and police chiefs. For the sake of consistency of our methodology, though, we do not manually separate these clusters. This cluster comprises around 5% of the work population and a more sophisticated industrial classification would likely help split these groups into two or more separated network communities.

Despite these limitations, our analysis provides a new way to reveal socioeconomic stratification based on the occupations and industries in which people work. Most research and policy approaches in economics on inclusive growth have been somewhat neutral about different types of occupations and industries. These approaches highlight the general need for (1) more education and human capital (Tavares and Menezes-Filho, 2011), (2) higher (minimum) wages for workers (López-Calva and Lustig 2010; Lustig et al., 2013), or (3) emphasize the need of promoting industrial growth and diversification at the macro level (Gala et al., 2017; Hartmann et al., 2017). These general measures may fail, though, if they do not consider the complex network structures of socioeconomic stratification and inequality that can hamper social cohesion and interactive learning. In contrast, approaches using large data sets and methods from network science are still not widespread in sociology. Here, we show how these methods can contribute to understanding the complex network structure of inequality and social stratification in modern societies.

Acknowledgements

DH acknowledges the financial support of the Inter-American Development Bank and the São Paulo Research Foundation (PROCESSO 2017/19842-2). CJF acknowledges support from The MIT Media Lab consortia. We also thank Elena Arias Ortiz, Mayra Bezerra, Robert Manduca, Benedikt Fritz, and participants from the Human Development Conference 2018 in Buenos Aires for valuable comments. Moreover, we are grateful Elton Freitas, A. Simoes, D. Landry, and Manuel Aristaran for data assistance.

6 References

Acemoglu, D. and Robinson, J. (2012) *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York: Crown.

Azar, J., Marinescu, I., & Steinbaum, M. I. (2017) *Labor market concentration*, National Bureau of Economic Research No. w24147.

Azar, J. A., Marinescu, I., Steinbaum, M. I., & Taska, B. (2018) *Concentration in US labor markets: Evidence from online vacancy data*, National Bureau of Economic Research No. w24395.

Alstott, J., Triulzi, G., Yan, B., & Luo, J. (2017) 'Mapping technology space by normalizing patent networks', *Scientometrics*, 110(1), 443-479.

Binder, A. J., Davis, D. B., & Bloom, N. (2016) 'Career funneling: How elite students learn to define and desire ''prestigious'' jobs', *Sociology of Education*, 89(1), 20-39.

Blau, P. M., & Duncan, O. D. (1967) *The American occupational struct*ure.

Blondel, V.D., Guillaume, J.L., Lambiotte, R. and Lefebvre, E., (2008) '*Fast unfolding of communities in large networks*', *Journal of statistical mechanics: theory and experiment*, 2008(10), 10008.

Bourdieu, P. (1984) *Distinction: A Social Critique of the Judgement of Taste*, translated by Richard Nice, NewYork, Cambridge University Press.

Bourdieu, P. (2011) [1986] 'The forms of capital', In Szeman, I. and Kaposy, T. (eds), *Cultural theory: An anthology*, Wiley-Blackwell, UK, 81-93.

Bourguignon, F., Ferreira, F.H. and Menéndez, M., (2007) 'Inequality of opportunity in Brazil', *Review of income and Wealth*, 53(4), 585-618.

Brummund, P., & Connolly, L. (2018) 'Who creates stable jobs? Evidence from Brazil', *Oxford Bulletin of Economics and Statistics*, 81(3), 540-563

Campbell, R. E. (1960) 'The prestige of industries', *Journal of Applied Psychology*, 44(1), 1.

Cardoso A, et al. (2007) *International Microdata Scoping Studies Project*, Economic and Social Research Council, Rio de Janeiro.

Clark, A. E., and Senik, C. (2010) 'Who compares to whom? The anatomy of income comparisons in Europe', *The Economic Journal*, 120(544), 573-594.

Coleman, J. S. (2000) 'Social capital in the creation of human capital'. In Lesser, E. (ed), *Knowledge and Social capital: Foundations and Applications*, Boston, Butterworth, 17-41.

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979) 'Intergenerational class mobility in three Western European societies: England, France and Sweden', *The British Journal of Sociology*, 30(4), 415-441.

Engerman, Stanley L., and Kenneth L. Sokoloff. (1997) 'Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies'. In Haber, S.H. (ed) *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914*, California, Stanford University Press, 260–304

Fortunato, S. and Hric, D. (2016) 'Community detection in networks: A user guide', *Physics Reports*, 659, 1-44.

Freyre, G. (1933) *Casa grande e senzala*, Rio de Janeiro, Schmidt Editor.

Fujiwara, T., Laudares, H., and Caicedo, F. V. (2017) *Tordesillas, Slavery and the Origins of Brazilian Inequality*.

Furtado, C. (1959) *Formação Econômica Do Brasil*, Rio de Janeiro, Fundo de Cultura

Furtado, C. (2009) O Nordeste e a Saga da Sudene. Rio de Janeiro, Contraponto

Gala, P., Camargo, J., & Freitas, E. (2017) 'The Economic Commission for Latin America and the Caribbean (ECLAC) was right: scale-free complex networks and core-periphery patterns in world trade' *Cambridge Journal of Economics*, 42(3), 633-651.

Gala, P., Camargo, J., Magacho, G., & Rocha, I. (2018) ‘Sophisticated jobs matter for economic complexity: An empirical analysis based on input-output matrices and employment data’, *Structural Change and Economic Dynamics*, 45, 1-8.
Grusky, D. B., & Sørensen, J. B. (1998) ‘Can class analysis be salvaged?’, *American Journal of Sociology*, 103(5), 1187-1234.
Guevara, M. R., Hartmann, D., & Mendoza, M. (2016) ‘diverse: an R package to analyze diversity in complex systems’, *R Journal*, 8(2), 60-78.
Guven, C., & Sørensen, B. E. (2012) ‘Subjective well-being: Keeping up with the perception of the Joneses’, *Social Indicators Research*, 109(3), 439-469.
Hartmann, D. (2014) *Economic Complexity and Human Development: How Economic Diversification and Social Networks Affect Human Agency and Welfare*. New York, Routledge.
Hartmann, D., Bezerra, M., Lodolo, B., & Pinheiro, F. L. (2019) *International Trade, Development Traps, and the Core-Periphery Structure of Income Inequality*, SSRN No. 3312097
Hartmann, D., Bezerra, M., & Pinheiro, F. L. (2019) *Identifying smart strategies for economic diversification and inclusive growth in developing economies*, SSRN No. 3346790
Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., and Hidalgo, C. A. (2017) ‘Linking economic complexity, institutions, and income inequality’, World Development, 93, 75-93.
Hartmann, D., Jara-Figueroa, C., Guevara, M., Simoes, A. and Hidalgo, C.A. (2016) ‘The structural constraints of income inequality in Latin America’, *Integration and Trade Journal*, 40, 70-85
Heckman, J.J. and Mosso, S. (2014) ‘The economics of human development and social mobility’, *Annu. Rev. Econ*., 6(1), 689-733.
Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann. (2007) ‘The Product Space Conditions the Development of Nations’, *Science*, 317 (5837), 482–87.
Hidalgo, C. A., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., ... & Neffke, F. (2018) ‘The principle of relatedness’, *International Conference on Complex Systems*, Springer, Cham, 451-457.
Jara-Figueroa, C., Jun, B., Glaeser, E. L., & Hidalgo, C. A. (2018) ‘The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms’, *Proceedings of the National Academy of Sciences*, *115*(50), 12646-12653.
Kalleberg, A. L. and Berg, I. (1987) *Work and industry: Structures, markets, and processes*, Springer Science & Business Media.
Kaltenberg and Hidalgo, (2019) *The Knowledge Manager: Wage Premiums in Knowledge Diverse Industries*.
Krueger, A. B., & Summers, L. H. (1988) ‘Efficiency wages and the inter-industry wage structure’, *Econometrica*, 259-293.
Kim, C. and Sakamoto, A. (2008) ‘The rise of intra-occupational wage inequality in the United States, 1983 to 2002’, American Sociological Review, 73(1), pp.129-157.
Lin, N. (2017) ‘Building a network theory of social capital’, In Lin, N., Cook, K., Burt, R.S. (eds) *Social Capital Theory and Research*, UK, Routledge, 3-28.
López-Calva, L. F., and Lustig, N. (2010) *Declining Inequality in Latin America: A Decade of Progress?*, Brookings Inst Press.
Lustig, N., Lopez-Calva, L.F. and Ortiz-Juarez, E., (2013) ‘Declining inequality in Latin America in the 2000s: the cases of Argentina, Brazil, and Mexico’, *World Development*, 44, 129-141.

Luttmer, E. F. (2005) 'Neighbors as negatives: Relative earnings and well-being', *The Quarterly Journal of Economics*, 120(3), 963-1002.
Maia, A.G. and Sakamoto, A., (2015) 'Occupational structure and socioeconomic inequality: a comparative study between Brazil and the United States', *Economia e Sociedade*, 24(2), 229-261.
Marx, K. (1867) *Das Kapital: Kritik der politischen Ökonomie, Vol. 1*, Hamburg, Germany, Otto Meisner.
Mazzon, J. A., & Kamakura, W. A. (2016) *Estratificação socioeconômica e consumo no Brasil*, Editora Blucher.
Ministério do Trabalho e Emprego (2010) *Códigos, títulos e descrições. Classificação Brasileira de Ocupações*, Brasília, 3ª edição.
Mouw, T., and Kalleberg, A. L. (2010) 'Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s', *American Sociological Review*, 75(3), 402-431.
Norton, M. I., and Ariely, D. (2011) 'Building a better America—One wealth quintile at a time', Perspectives on psychological science, 6(1), 9-12.
Novotný, J. (2007) 'On the measurement of regional inequality: does spatial dimension of income inequality matter?', *The Annals of Regional Science*, 41(3), 563-580.
Peixoto, T. P. (2014) 'Efficient Monte Carlo and greedy heuristic for the inference of stochastic block models', *Physical Review E*, 89(1), 012804.
Prebisch, R. (1949) *The Economic Development of Latin America and Its Principal Problems*, United Nations.
Rinz, K. (2018) *Labor market concentration, earnings inequality, and earnings mobility*, Center for Administrative Records Research and Applications Working Paper 10.
Rodríguez-Castelán, C. and López-Calva, L. F. and Lustig, N. and Valderrama, D. (2016) *Understanding the Dynamics of Labor Income Inequality in Latin America*, World Bank Policy Research Working Paper No. 7795.
Rodrik, D. and McMillan, M. (2011) *Globalization, Structural Change and Productivity Growth*, NBER Working Paper No. 17143
Rodrik, D. (2016) 'Premature deindustrialization', *Journal of Economic Growth*, 21(1), 1-33.
Rytina, S. (1992) 'Scaling the intergenerational continuity of occupation: Is occupational inheritance ascriptive after all?', *American Journal of Sociology*, 97(6), 1658-1688.
Tatro, C. R., & Garbin, A. P. (1973) 'The industrial prestige hierarchy', *Journal of Vocational Behavior*, 3(3), 383-391.
Tavares, P. A., & Menezes-Filho, N. A. (2011) 'Human capital and the recent fall of earnings inequality in Brazil', *Brazilian review of econometrics*, 31(2), 231-257.
Weber, M. (1922) *Wirtschaft und Gesellschaft*, Tübingen, Mohr
Weeden, K. A., and Grusky, D. B. (2005) 'The case for a new class map', *American Journal of Sociology*, 111(1), 141-212.
Weeden, K. A., & Grusky, D. B. (2012) 'The three worlds of inequality', *American Journal of Sociology*, 117(6), 1723-1785.
Wilmers, N. (2018) 'Wage stagnation and buyer power: How buyer-supplier relations affect US workers' wages, 1978 to 2014', *American Sociological Review*, 83(2), 213-242.
Xie, Y., Killewald, A. and Near, C., (2016) 'Between-and within-occupation inequality: The case of high-status professions', *The Annals of the American Academy of Political and Social Science*, 663(1), pp.53-79.
Zhou, X., & Wodtke, G. T. (2018) 'Income Stratification among Occupational Classes in the United States', *Social Forces*, 97(3), 945-972.

7 Appendix

Figure S1. A. Neighborhood connectivity, and B. average path lengths, of the occupations in the BIOS

BIOS Occupation grouping	Size BIOS occupation grouping	CBO occupation grouping	Size CBO occupation grouping	Adjusted Rand Index (ARI)
Peixoto& Louvain	**28**	**ocode3**	**187**	**0.093**
Peixoto& Louvain	**28**	**ocode2**	**45**	**0.195**
Peixoto& Louvain	**28**	**ocode1**	**9**	**0.103**
Louvain	21	ocode3	187	0.072
Louvain	21	ocode2	45	0.191
Louvain	21	ocode1	9	0.121
Peixoto	12	ocode3	187	0.010
Peixoto	12	ocode2	45	0.032
Peixoto	12	ocode1	9	0.025

Table S1. Adjusted Rand Indices (ARI) comparing BIOS network cluster solutions with different aggregation levels of the Brazilian Occupational Classification (CBO).

10 occupations with highest share of women

ID	Community	Occupation name	f_male
2238	18	Audiologists	0.045
2237	18	Nutritionists	0.057
3311	19	Early Childhood Teachers	0.063
2516	18	Social Workers	0.070
2392	19	Special Education Teachers	0.083
2239	18	Physical Therapists	0.083
5162	19	Caregivers	0.086
2311	19	Elementary School Teachers	0.099
5133	0	Housekeepers	0.099
3224	18	Dental Technicians	0.116

10 occupations with highest share of men

ID	Community	Occupation name	f_male
7153	1	Reinforced Concrete Assemblers	0.993
6420	9	Forest Mechanization Workers	0.993
9131	9	Heavy Machinery Mechanics	0.992
7155	1	Construction Workers	0.992
7152	1	Masonry Workers	0.992
7825	9	General Cargo Drivers	0.992
7113	9	Liquid and Gas Extraction Workers	0.991
7154	1	Concrete Machine Operators	0.991
7151	9	Excavators	0.989
7824	13	Public Transit Driver	0.989

Table S2. Ten occupations with highest shares of women or men

10 occupations with highest level of race diversity

ID	Community	Occupation name	hh
6323	9	Fiber, Wax, and Oil Extraction Workers	0.330
7113	9	Liquid and Gas Extraction Workers	0.348
7114	5	Salt Workers	0.353
8412	0	Salt Processing Workers	0.354
3522	18	Health and Environmental Workers	0.354
7832	9	Movers	0.355
8113	23	Filtration Workers	0.355
2541	0	Federal Revenue Inspectors	0.358
8486	0	Craft Tobacoo Workers	0.358
3161	9	Geology Workers	0.359

10 occupations with lowest race diversity

ID	Community	Occupation name	hh
2153	9	Pilots	0.759
2331	19	Vocational Teachers	0.749
2033	6	Health Science Researchers	0.745
2122	17	Computer Engineers	0.731
7256	13	Aircraft Assemblers	0.730
2542	0	Welfare Inspectors	0.729
1113	17	Magistrates	0.725
2031	23	Natural Science Researchers	0.707
1417	17	Financial Institution Managers	0.701
2146	25	Material Science Engineers	0.700

Table S3. Ten occupations with highest and lowest shares of race diversity

Hohenheim Discussion Papers in Business, Economics and Social Sciences

This paper series aims to present working results of researchers of all disciplines from the Faculty of Business, Economics and Social Sciences and their cooperation partners since 2015.

Institutes

510	Institute of Financial Management
520	Institute of Economics
530	Institute of Health Care & Public Management
540	Institute of Communication Science
550	Institute of Law and Legal Sciences
560	Institute of Education, Labour and Society
570	Institute of Marketing & Management
580	Institute of Interorganizational Management & Performance

Research Areas (since 2017)

INEPA	"Inequality and Economic Policy Analysis"
TKID	"Transformation der Kommunikation – Integration und Desintegration"
NegoTrans	"Negotiation Research – Transformation, Technology, Media and Costs"
INEF	"Innovation, Entrepreneurship and Finance"

The following table shows recent issues of the series. A complete list of all issues and full texts are available on our homepage: https://wiso.uni-hohenheim.de/papers

No.	Author	Title	Inst
01-2018	Michael D. Howard Johannes Kolb	FOUNDER CEOS AND NEW VENTURE MEDIA COVERAGE	INEF
02-2018	Peter Spahn	UNCONVENTIONAL VIEWS ON INFLATION CONTRAOL: FORWARD GUIDANCE, THE NEO-FISHERIAN APPROACH, AND THE FISCAL THEORY OF THE PRICE LEVEL	520
03-2018	Aderonke Osikominu Gregor Pfeifer	PERCEIVED WAGES AND THE GENDER GAP IN STEM FIELDS	INEPA
04-2018	Theresa Grafeneder-Weissteiner Klaus Prettner Jens Südekum	THREE PILLARS OF URBANIZATION: MIGRATION, AGING, AND GROWTH	INEPA
05-2018	Vadim Kufenko Vincent Geloso Klaus Prettner	DOES SIZE MATTER? IMPLICATIONS OF HOUSEHOLD SIZE FOR ECONOMIC GROWTH AND CONVERGENCE	INEPA
06-2018	Michael Trost	THE WHOLE IS GREATER THAN THE SUM OF ITS PARTS – PRICING PRESSURE INDICES FOR MERGERS OF VERTICALLY INTEGRATED FIRMS	520
07-2018	Karsten Schweikert	TESTING FOR COINTEGRATION WITH TRESHOLD ADJUSTMENT IN THE PRESENCE OF STRUCTURAL BREAKS	520
08-2018	Evanthia Fasoula Karsten Schweikert	PRICE REGULATIONS AND PRICE ADJUSTMENT DYNAMICS: EVIDENCE FROM THE AUSTRIAN RETAIL FUEL MARKET	520

No.	Author	Title	Inst
09-2018	Michael Ahlheim Jan Neidhardt Ute Siepmann Xiaomin Yu	WECHAT – USING SOCIAL MEDIA FOR THE ASSESSMENT OF TOURIST PREFERENCES FOR ENVIRONMENTAL IMPROVEMENTS IN CHINA	520
10-2018	Alexander Gerybadze Simone Wiesenauer	THE INTERNATIONAL SALES ACCELERATOR: A PROJECT MANAGEMENT TOOL FOR IMPROVING SALES PERFORMANCE IN FOREIGN TARGET MARKETS	570
11-2018	Klaus Prettner Niels Geiger Johannes Schwarzer	DIE WIRTSCHAFTLICHEN FOLGEN DER AUTOMATISIERUNG	INEPA
12-2018	Martyna Marczak Thomas Beissinger	COMPETITIVENESS AT THE COUNTRY-SECTOR LEVEL: NEW MEASURES BASED ON GLOBAL VALUE CHAINS	520
13-2018	Niels Geiger Klaus Prettner Johannes Schwarzer	AUTOMATISIERUNG, WACHSTUM UND UNGLEICHHEIT	INEPA
14-2018	Klaus Prettner Sebastian Seiffert	THE SIZE OF THE MIDDLE CLASS AND EDUCATIONAL OUTCOMES: THEORY AND EVIDENCE FROM THE INDIAN SUBCONTINENT	INEPA
15-2018	Marina Töpfer	THE EFFECT OF WOMEN DIRECTORS ON INNOVATION ACTIVITY AND PERFORMANCE OF CORPORATE FIRMS - EVIDENCE FROM CHINA –	INEF
16-2018	Timo Walter	TRADE AND WELFARE EFFECTS OF A POTENTIAL FREE TRADE AGREEMENT BETWEEN JAPAN AND THE UNITED STATES	INEPA
17-2018	Jonas Frank	THE EFFECTS OF ECONOMIC SANCTIONS ON TRADE: NEW EVIDENCE FROM A PANEL PPML GRAVITY APPROACH	INEPA
18-2018	Jonas Frank	THE EFFECT OF CULTURE ON TRADE OVER TIME – NEW EVIDENCE FROM THE GLOBE DATA SET	520
19-2018	Dario Cords Klaus Prettner	TECHNOLOGICAL UNEMPLOYMENT REVISITED: AUTOMATION IN A SEARCH AND MATCHING FRAMEWORK	INEPA
20-2018	Sibylle Lehmann-Hasemeyer Andreas Neumayer	THE PERSISTENCE OF OWNERSHIP INEQUALITY – INVESTORS ON THE GERMAN STOCK EXCHANGES, 1869-1945	INEPA
21-2018	Nadja Dwenger Lukas Treber	SHAMING FOR TAX ENFORCEMENT: EVIDENCE FROM A NEW POLICY	520
22-2018	Octavio Escobar Henning Mühlen	THE ROLE OF FDI IN STRUCTURAL CHANGE: EVIDENCE FROM MEXICO	520

No.	Author	Title	Inst
24-2018	Peng Nie Lanlin Ding Alfonso Sousa-Poza	OBESITY INEQUALITY AND THE CHANGING SHAPE OF THE BODYWEIGHT DISTRIBUTION IN CHINA	INEPA
25-2018	Michael Ahlheim Maike Becker Yeniley Allegue Losada Heike Trastl	WASTED! RESOURCE RECOVERY AND WASTE MANAGEMENT IN CUBA	520
26-2018	Peter Spahn	WAS WAR FALSCH AM MERKANTILISMUS?	520
27-2018	Sophie Therese Schneider	NORTH_SOUTH TRADE AGREEMENTS AND THE QUALITY OF INSTITUTIONS: PANEL DATA EVIDENCE	INEPA
01-2019	Dominik Hartmann Mayra Bezerra Beatrice Lodolo Flávio L. Pinheiro	INTERNATIONAL TRADE, DEVELOPMENT TRAPS, AND THE CORE-PERIPHERY STRUCTURE OF INCOME INEQUALITY	INEPA
02-2019	Sebastian Seiffert	GO EAST: ON THE IMPACT OF THE TRANSIBERIAN RAILWAY ON ECONOMIC DEVELOPMENT IN EASTERN RUSSIA	INEPA
03-2019	Kristina Bogner	KNOWLEDGE NETWORKS IN THE GERMAN BIOECONOMY: NETWORK STRUCTURE OF PUBLICLY FUNDED R&D NETWORKS	520
04-2019	Dominik Hartmann Mayra Bezerra Flávio L. Pinheiro	IDENTIFYING SMART STRATEGIES FOR ECONOMIC DIVERSIFICATION AND INCLUSIVE GROWTH IN DEVELOPING ECONOMIES. THE CASE OF PARAGUAY	INEPA
05-2019	Octavio Escobar Henning Mühlen	DECOMPOSING A DECOMPOSITION: WITHIN-COUNTRY DIFFERENCES AND THE ROLE OF STRUCTURAL CHANGE IN PRODUCTIVITY GROWTH	INEPA
06-2019	Dominik Hartmann Cristian Figueroa Mary Kaltenberg Paolo Gala	MAPPING STRATIFICATION: THE INDUSTRY-OCCUPATION SPACE REVEALS THE NETWORK STRUCTURE OF INEQUALITY	INEPA

IMPRINT

University of Hohenheim
Dean’s Office of the Faculty of Business, Economics and Social Sciences
Palace Hohenheim 1 B
70593 Stuttgart | Germany
Fon +49 (0)711 459 22488
Fax +49 (0)711 459 22785
wiso@uni-hohenheim.de
wiso.uni-hohenheim.de