Specificity of Human Capital: An Occupation Space Based on Job-to-Job Transitions

Eduardo Levy Yeyati and Martín Montané

CID Faculty Working Paper No. 379
April 2020
Revised May 2020

Working Papers
Center for International Development
at Harvard University

Specificity of human capital:

Occupation and industry spaces

based on job-to-job transitions

Eduardo Levy Yeyati

Martín Montané [1] [2]

Abstract

Using job transition data from Argentina's Household Survey, we document the extent to which human capital is specific to occupations and activities. Based on workers' propensity to move between occupations/industries, we build Occupation and Industry Spaces to illustrate job similarities, and we compute an occupation and industry similarity measures that, in turn, we use to explain wage transition dynamics. We show that our similarity measures influence positively post-transition wages. Inasmuch as wages capture a worker´s marginal productivity and this productivity reflects the degree to which a worker matches the job's skill demand, our results indicate that a worker´s human capital is specific to both occupation and activity: closer occupations share similar skill demands and task composition (in other words, demand similar workers) and imply a smaller human capital loss in the event of a transition.

[1] Eduardo Levy Yeyati is the Dean of the School of Government, Universidad Torcuato Di Tella; Martín Montané is the director of the Data Lab of the School of Government, Universidad Torcuato Di Tella. The usual disclaimers apply.

[2] We believe that replication of scientific studies is a necessary condition for the improvement of science. Codes and data are publicly available at https://github.com/martinmontane/OcupacionesEPH_WP. Any suggestion about the code or the data will be welcome, please direct them to martinmontane@gmail.com.

Introduction

The question of whether human capital is general or specific has been extensively studied in recent decades. Two factors have contributed strongly to this phenomenon: 1) a growing penetration of technology that could replace labor in various production processes, and the concerns associated with technological displacement, and 2) the emergence of new high-quality datasets relevant to these investigations.

Advances on this topic have appeared mainly in countries where administrative data is most accessible or long-standing panel type surveys are available. Because of that, the literature that analyzes labor mobility in a more granular form are available only for developed economies such as the United States, Germany or the UK, with conclusions that cannot be extended to developing economies without accounting for differences in human capital and labor regulations, and the prevalence of informality. For example, there is evidence that the difference between the skills demands in jobs before and after the transition correlates with the size of wage losses of displaced workers in the US (Poletaev and Robinson, 2008). Also, there is evidence that human capital is specific to occupations (in the US: Kambourov and Manovskii, 2009, and in the UK: Zangelidis, 2008), to tasks (in Germany: Gathmann & Schonberg, 2010), or to both occupation and industry (again in the US: Sulivan, 2010).

In this paper, we build on the methodology used in Neffke, F., Otto and Weyh (2017) and address the question of the specificity of human capital for the case of Argentina –to our knowledge, this is the first such analysis reported for a developing economy.[3]

We work in two steps. First, we construct Occupation and Industry Spaces based on job turnover data –specifically, on the propensity to move from one occupation/industry to the rest– from which we extract a measure of the relative similarity between pairs of occupations and industries. Second, we estimate the incidence of this similarity measure on the new wages when workers change occupations/industries, and show that similarity helps predict the new wage after a change. We interpret this result as indicating that the relative similarity in the Occupation and Industry Spaces captures, at least partially, the nature of the worker´s specific human capital: closer occupations share similar skill demands and task composition –in other words, demand similar workers– under the assumption that the wage measures the worker´s marginal productivity and that this productivity is a function of the degree to which the worker matches the demands of the job.

---

[3] While our results are country-specific, the methodology is not, and be could easily extended to other countries, and to more granular industry/occupational/task data, whenever available.

To the extent that this relative similarity provides information about the characteristics (for example, the required skills, tasks and experience) of actual jobs, the findings of this paper could be useful to orient employment and reskilling services to occupations that are not “distant” from the worker´s experience –an aspect that may help mitigate the capital loss associated with job displacement and reallocation process. Similarly, similarity measures could help find new talent for growing industries facing a shortage in specific labor supply, to the extent to which hiring from similar occupations may minimize training costs.

Similarity between occupations: Methodology

The methodology used in this work is based on the theoretical contributions of the literature of economic complexity (Hidalgo et al., 2007; Hidalgo and Hausmann, 2009; Neffke and Henning, 2013; and Neffke et al. 2017). The objective of this literature was to identify similarities (for example, between national productive capacities, or skill demands) base on observed flows (international trade, workers´ transitions) between nodes (countries, occupations).

This approach requires a "base case" flow between the nodes, defined as the expected flow in the absence of a specific attraction between a pair of nodes. Once the expected flow is identified, this value is compared to the observed flow and conclusions are drawn based on the flow differences.

This way of detecting non-random patterns has analogies in other domains; for example, in spatial statistics, to compare crime rates in different regions relative to the average, which indicates a random distribution in spatial terms, or in the methodology used for contingency table tests (de Raco and Semeshenko, 2019).

However, the strategy used to determine the selection of this base scenario has varied throughout literature. Hidalgo et al. (2007) create this base scenario using Balassa's (1986) concept of Revealed Comparative Advantage (RCA). Analyzing international trade between countries they define the RCA of country *c* in product *p* as:

$$RCA_{cp} = (VE_{cp} / \sum_{p=1}^{P} VE_{cp}) \,/\, (\sum_{c=1}^{C} VE_{cp} \,/\, \sum_{c=1}^{C} \sum_{p=1}^{P} VE_{cp})$$

where VE_{cp} is the exported value of the product *p* by country *c*. Thus, a value less than one in this indicator implies that the country exports a lower proportion of that good relative to the global average, and a value greater than one implies the opposite. Therefore, the expected flow is the value that would have been observed if the export of that product were distributed evenly across countries.

This notion can be adapted to the movement of workers between occupations simply by observing the flows of workers (renamed *F*) coming out of occupation/industry *i* to occupation/industry *j*.

Following Neffke et al. (2017), we establish the expected flow of workers between occupations/industries combining the outflows from and inflows to the occupations/industries between which the workers moved. If the relative output and input size could be used to explain the flow of workers, then the expected flow between occupation/industry *i* and *j* would be given by $\widehat{F_{i,j}} = \frac{F_{i,*} * F_{*,j}}{F_{*,*}}$. In this way, we can measure the similarity between occupations/industries by the ratio between expected and observed flows, which, after a small algebraic arrangement, takes the following form:

$$R_{i,j} = \frac{F_{ij} * F_{**}}{F_{i,*} * F_{*,j}}$$

where asterisks (*) refer to all sectors. In this way, the similarity R_{ij} between occupation/industry *i* and occupation/industry *j* is given by a variable that goes between zero and infinite. A value smaller than 1 implies a movement of workers below the expectations, while one greater than 1 implies the opposite. However, this similarity has an asymmetric distribution, with extreme values on the positive side. To count with a measure that is symmetrically distributed around a value (0) we make the following transformation:

$$\overline{R_{i,j}} = \frac{(R_{ij} - 1)}{(R_{ij} + 1)}$$

This measure of similarity is now bounded between -1 and 1. Although the distance is now distributed between -1 and 1, the similarity measure is it not symmetrical in the sense that the distance between R_{ij} does not have to be equal to R_{ji}. This is not a problem for the regression analysis, since we exploit this asymmetry, but the graph analysis is simplified by taking the maximum value between each pair of occupations.

Data

We use data from Argentina´s Household Survey (EPH), which systematically collects information on the socio-economic situation of most of Argentina's urban areas, reaching approximately 64% of the country's population. The survey is carried out continuously and on a quarterly basis since the second quarter of 2003, with some missing quarters in 2007, 2015 and 2016. Like many household surveys in other countries, the sample design has a rotating nature: the same home is surveyed for four noncontiguous quarters. A household is surveyed twice in a row, stays out two quarters, and is surveyed again for two additional quarters. In this way, the job dynamics within an individual household is observed for a maximum time frame of approximately 18 months. This work exploits these time windows to work with short panels and compute transitions across occupations and industries and infer similarities between occupations/industries based on those transition flows. Although these are short trajectories, they are comparable to those used by Nedelkoska et al (2018), who looked at monthly data from the U.S. Current Population Survey (CPS), with a 15-month time window virtually identical to the one of the EPH.

Occupations are included in the National Occupational Code (CNO/2001). It is a hierarchical occupation classifier, in which the first two digits correspond to occupations with a high level of aggregation, while the rest of the digits refer to the level of qualification, technology, and hierarchy. Due to the sparse data, we work with the two-digits codes of the CNO, which refers to the occupational groups. In total, there are 52 large occupational groups (see Annex Table 1).

On the other hand, industries are classified by the Sociodemographic Economic Industry's Classification (CAES in Spanish). We also use two-digit codes given the low frequency in most of the three-digit codes. We show the two-digit CAES codes and their description in Table 1 in the Annex.

A factor relevant in the measurement of the similarity between occupations is the regional dimension of labor markets, namely, the presence of commuting zones. To the extent that flows can be in part explained by the simple fact that some occupations/industries are more in demand in some commuting zones, aggregated flows at the national level would mask this effect and measure similarity inconsistently. For this reason, in this paper we work only with data from the Buenos Aires metropolitan area, which for practical purposes can be considered as a single commuting zone. This leaves us with 201,587 individuals that were either employed, unemployed or in inactivity during the second quarter of 2003 and the second quarter of 2019.

Descriptive transition analysis

The first step to analyze the transition of workers is to determine where an actual job transition has occurred in the data. This is far from trivial in the case of survey data, as the literature has already noted (see for example Kambourov and Manovskii, 2009)[4]. In this paper, we use two rules to determine when an actual job transition has been made: a *seniority rule* and an *interim rule*.

The *seniority rule* is an adapted version of the *Partition T* method of Brown and Light (1992) or Kambourov and Manovskii (2009). In the *Partition T* method, a transition is detected whenever the elapsed time between two observations of the same individual is greater than the job tenure that

[4] For displaced workers data the situation is obviously easier, although there is a clear bias in the sample: displaced workers may not transition in the same fashion as the rest of the workers.

should have been reported in the second observation if the individual had continued working in the job reported in the first observation. Given the short time frame and the fact that we have bracketed information on job tenure, we detect a job transition under the seniority rule if the reported experience in the job is greater than the time elapsed since an individual entered the sample. In the Annex 2 we explain the variables that were used and the specific brackets that the EPH surveys each quarter.

On the other hand, the *interim rule* detects a job transition if there is a quarter of unemployment or inactivity between two non-contiguous quarters of reported employment. Based on these conditions, we identify 10,371 job changes between 2003 and 2019.

Before displaying the analysis of the resulting graph and estimating the models, in this section we focus on the detected transitions and the robustness of the switch detection. Of the 10,371 transitions, 3,717 correspond to the *interim rule*, while the remaining 6,654 were detected through the seniority rule.

The *interim rule* is robust since there is no way of modifying the criteria: if the individual reported being unemployed or inactive during a period between two observations of employment, we say that there was a job switch. The *seniority rule*, on the other hand, is subject to more arbitrariness in the criteria. We opted to take the maximum possible value in the reported bracket as to minimize the possibility of detecting too many job switches, which would magnify the detected job switches in the same occupational and industry codes.

The importance of differentiating these labor transitions is that we expect a different wage impact: specifically, ceteris paribus, the wage change after an unemployment spell may be differ from that after a job-to-job transition. On the one hand, people may be inactive/unoccupied during the transition because they are more demanding, so that, when they finally get a job, it is on average better than those that move without waiting. On the other hand, it is also likely that those who go through unemployment or inactivity should be cash-strapped or may have more difficulty in getting a new job, including because of the negative signaling effect of unemployment, so that they may be willing to move to an occupation further away from its expertise and accept a lower pay.

From the point of view of the literature, it is also relevant to analyze both types of transitions since many of the papers are based on administrative data from displaced workers and it would be interesting to find out if there are differences between the two groups, sometimes a question overlooked in the relevant literature (see for example Poletaev and Robinson, 2008 or Gathmann & Schonberg, 2010)

Which are the most normal occupations and industry transitions for all the detected job transitions? Figure 1 shows the share of job-to-job transitions decomposed by “old occupation” and “new occupation”. Among the largest four occupations, which account for more than 50% of job-to-job transitions, direct commercialization “pulled” more workers than it “pushed” towards other occupations. It is not a surprise to find construction occupations as one of the most important occupational codes, given its relevance in the economy and the short-term nature of the employment relationship.

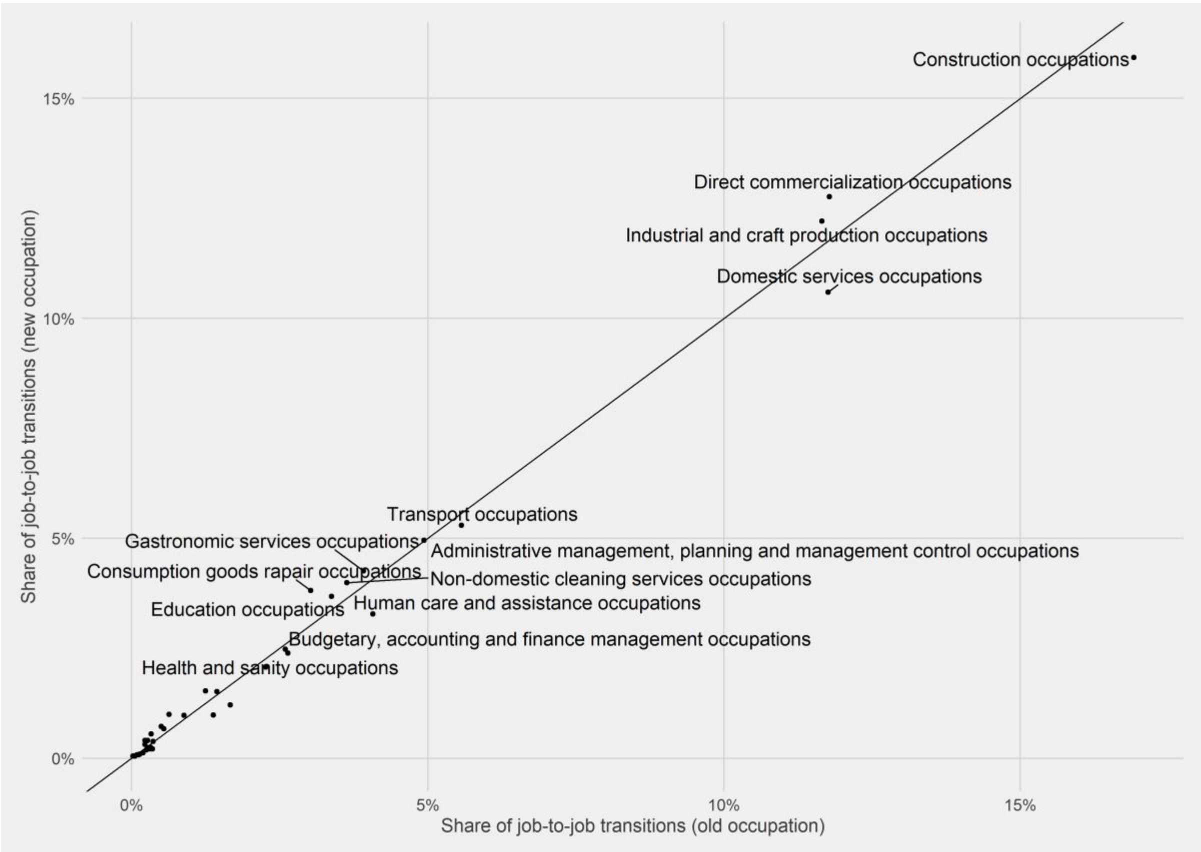

Figure 1. Distribution of occupations according to their participation in the total of transitions such as departure (old) and arrival (new) occupations. Elaborated by the authors based on the Permanent Household Survey (EPH).

When we turn to industry codes, we face a similar picture: construction, commerce and domestic services alone comprises almost 50% of all job-to-job switches. Once again, commerce seems to be attracting more workers than it pushes to other industry codes.

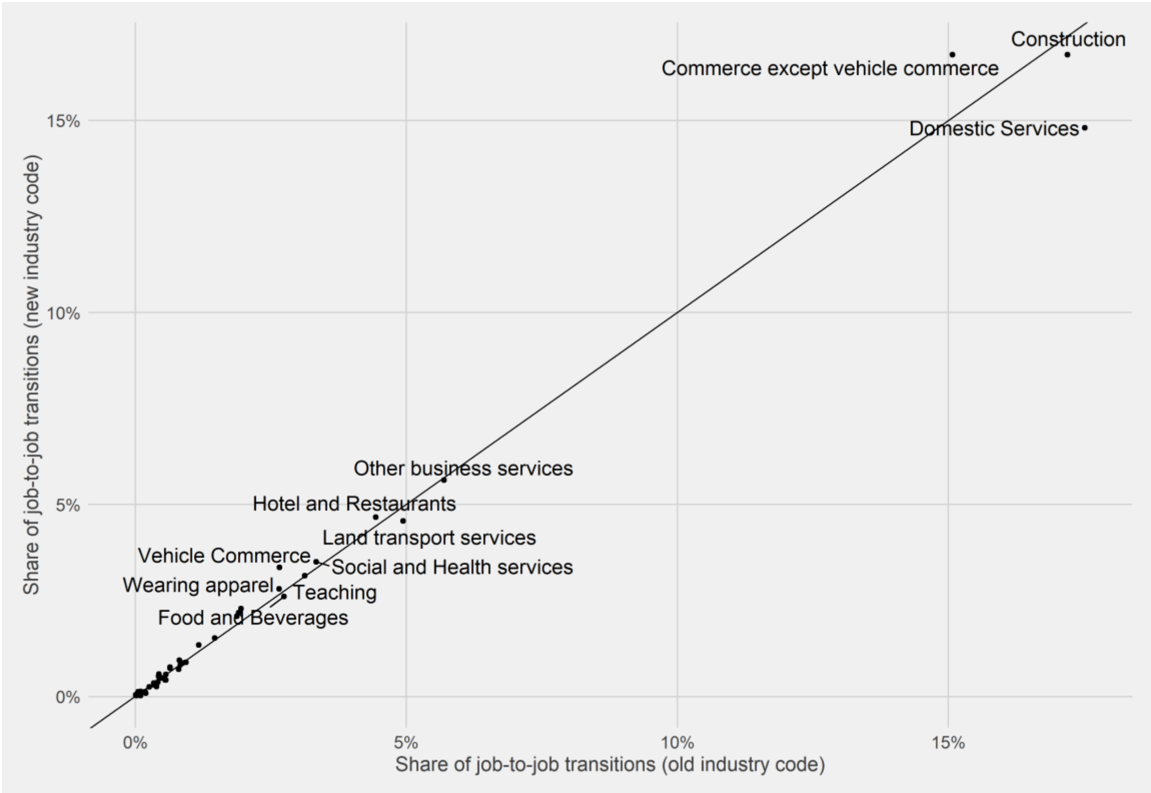

Figure 2. Distribution of occupations according to their participation in the total of transitions such as departure (old) and arrival (new) industry. Elaborated by the authors based on the Permanent Household Survey (EPH).

How does our similarity measure correlate with the age and educational level of the workers? We expect the older and more educated the worker, the more similar the occupations or activities they transition between –much in the same way as it has been reported for task similarity by displaced workers in Germany (Gathmann & Schonberg, 2010).

The intuition is simple: as workers accumulate more human capital, inasmuch as it is specific, it becomes costlier (in terms of human capital) to switch between distant occupations/industries. Workers with less human capital do not face the same incentives and therefore should be more willing to move to more dissimilar occupations or activities.

Our data is in line with this hypothesis: the younger the worker, the more dissimilar the transition. This is specially the case for workers of 30 years old or less, a stylized fact –both in the occupation and industry similarity measures (Figure 3)– that deserves more careful empirical examination. In the same vein, we find that workers with a university degree are significantly more likely than their less educated counterparts to switch between dissimilar occupations/industries (Figure 4).

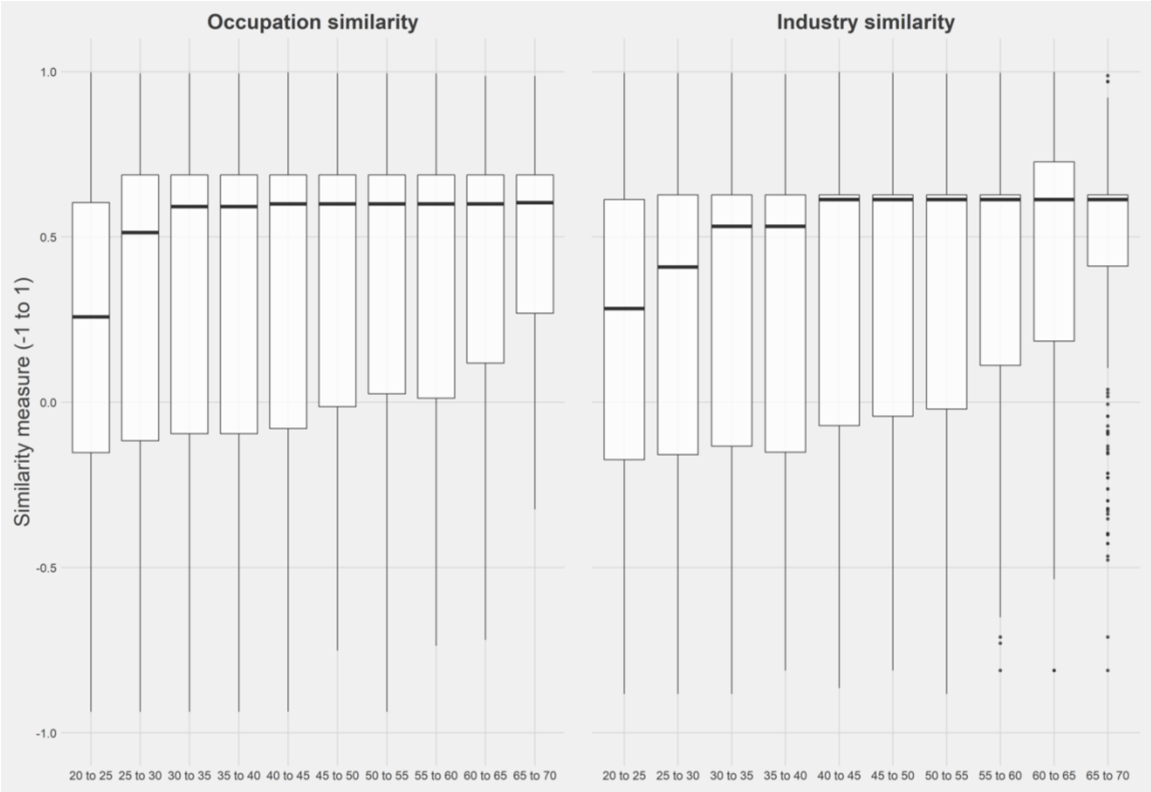

Figure 3. Boxplot of the distribution of occupational similarity (left size) and industries (right size) for different age groups. The limits of the age groups are stated in the x axis and are expressed in years.

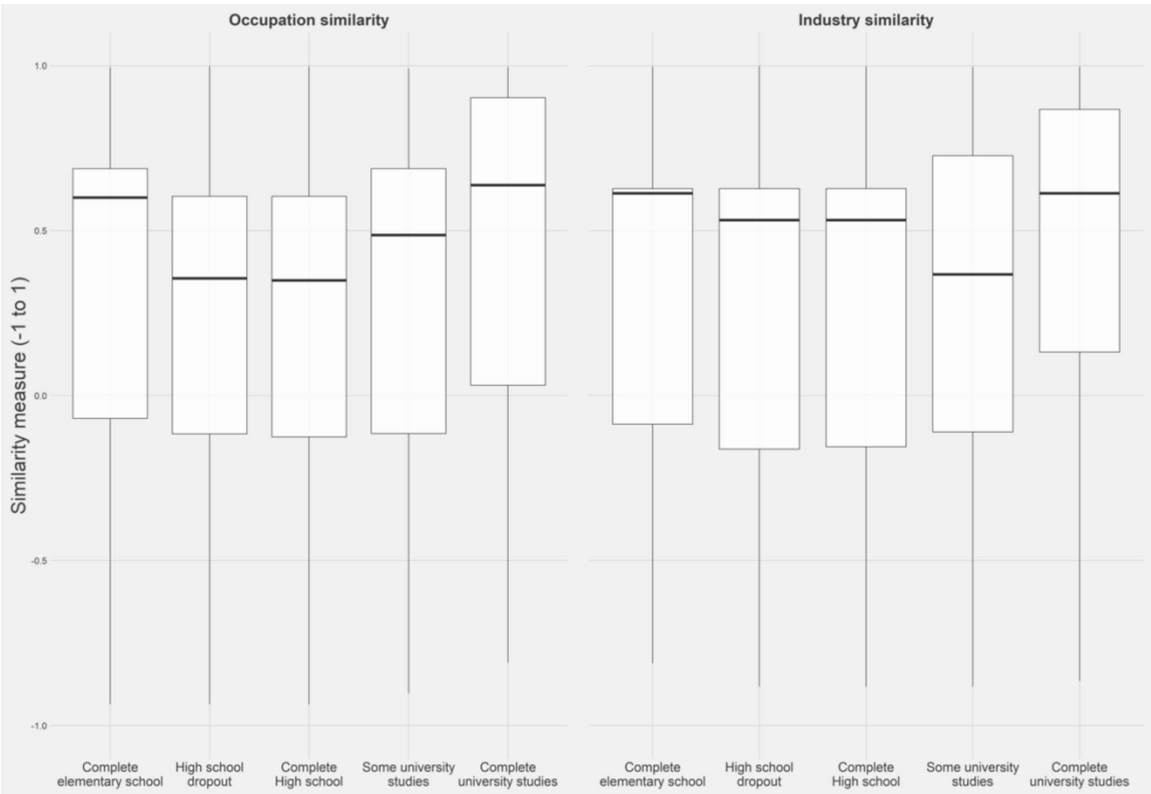

Figure 4. Boxplot of the distribution of the similarity measure of occupations (left size) and industries (right size) for different educational groups.

Finally, if we put both stylized facts together and estimate a regression on our similarity occupational and industry similarity measures, we find that both facts hold even after accounting for the fact that age and educational are related between each other (Table 1). Younger and less educated workers tend to switch between relatively dissimilar occupations relative to older and more educated workers.

Dependent variable:	Occupational similarity (1)	Industry similarity (2)
Age (years) *Base: 20 to 25*		
25 to 30	0.05 ***	0.03
	(0.02)	(0.02)
30 to 35	0.09 ***	0.05 **
	(0.02)	(0.02)
35 to 40	0.09 ***	0.07 ***
	(0.02)	(0.02)
40 to 45	0.10 ***	0.10 ***
	(0.02)	(0.02)
45 to 50	0.13 ***	0.11 ***
	(0.02)	(0.02)
50 to 55	0.13 ***	0.13 ***
	(0.02)	(0.02)
55 to 60	0.14 ***	0.18 ***
	(0.02)	(0.02)
60 to 65	0.16 ***	0.21 ***
	(0.03)	(0.03)
65 to 70	0.22 ***	0.21 ***
	(0.03)	(0.03)
Maximum educational level *Base: Complete elementary studies*		
High school dropout	-0.00	-0.01
	(0.01)	(0.01)
Complete High school	-0.02	-0.00
	(0.01)	(0.01)
Some university studies	0.02	0.01
	(0.02)	(0.02)
Complete university studies	0.16 ***	0.15 ***

	(0.02)	(0.02)
Intercept	0.24 ***	0.24 ***
	(0.01)	(0.01)
N	8743	8187
R2	0.03	0.03

*** p < 0.001; ** p < 0.01; * p < 0.05.

Table 1: Summary table of two regressions. (1) has the occupational similarity measure as dependent variable, while (2) has the industry similarity measure as dependent variable. Both regressions have the same regressors that are stated in the first column.

Occupation and Industry spaces

Given that we captured relationships between occupations and industries we can leverage on it and show a graph of the connection between each node. Before showing the networks, Figures 5 and 6 can give us great insights on the occupation and industry flows.

Figure 5 shows the density map of similarity between the occupations analyzed with the formula based on Neffke et al (2017), centered on 0 and with limits between -1 and 1. As can be seen, the main diagonal shows the highest values in the measure of similarity derived simply from observing the flows between occupations and comparing them against a theoretical expected flow.

Even more interesting is the fact that the similarity measure appears to have high values in areas close to the main diagonal, while values represented by dark colors (low) are in the regions farther away from this diagonal, on average.

This pattern suggests that the most common occupational transitions occur between occupational groups that generally have some similarity according to the CNO. This can be inferred form the high values along the main diagonal, which shows the similarity between the same CNO code and neighboring codes. The density also shows how many occupations are disconnected based on job-to-job transitions (similarity value of -1), but only a few are heavily connected occupational codes.

When we turn to the similarity between industries, we see the same pattern (Figure 6). The main diagonal and small clusters show the highest values in our similarity measure. One interesting feature of this heat map is the many high attraction clusters that can be found between digits ranging from 15 to 40. All of these codes are manufacturing industries, such as textile or automobile, which shows that our similarity measure is picking some of the similarity embedded in the industry classification itself.

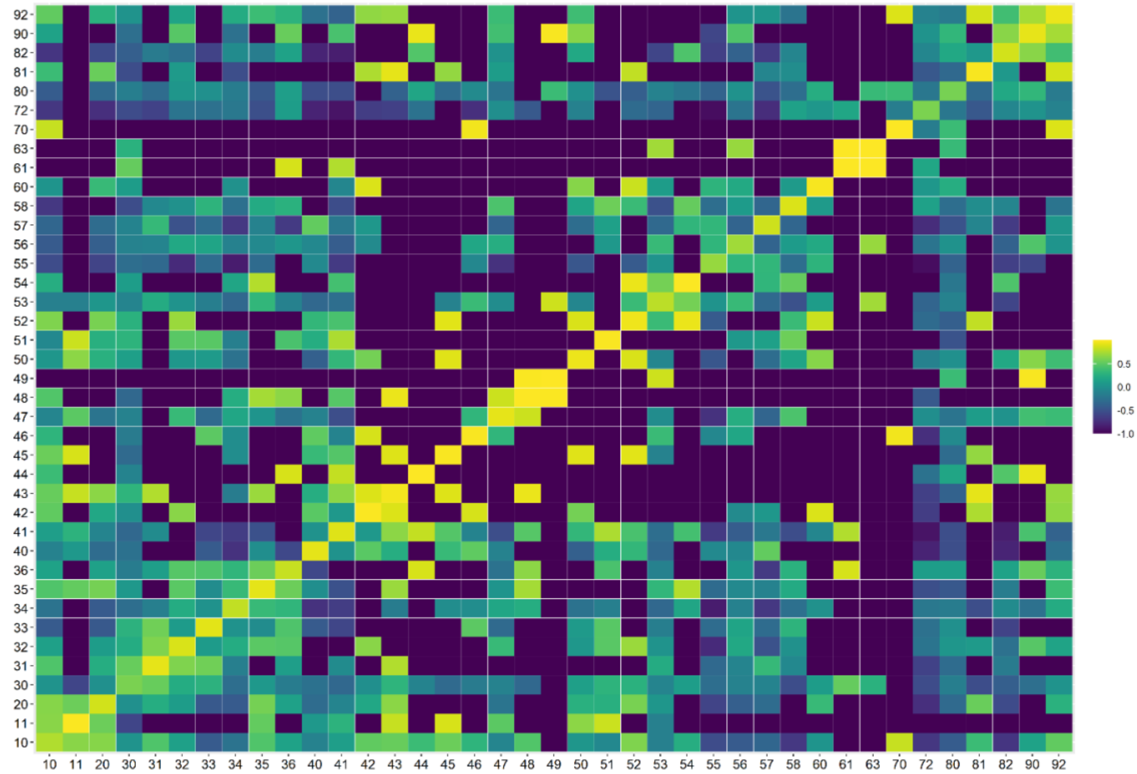

Figure 5. Density map among occupations of the National Occupation Code (CNO) based on data from the Permanent Household Survey (EPH)

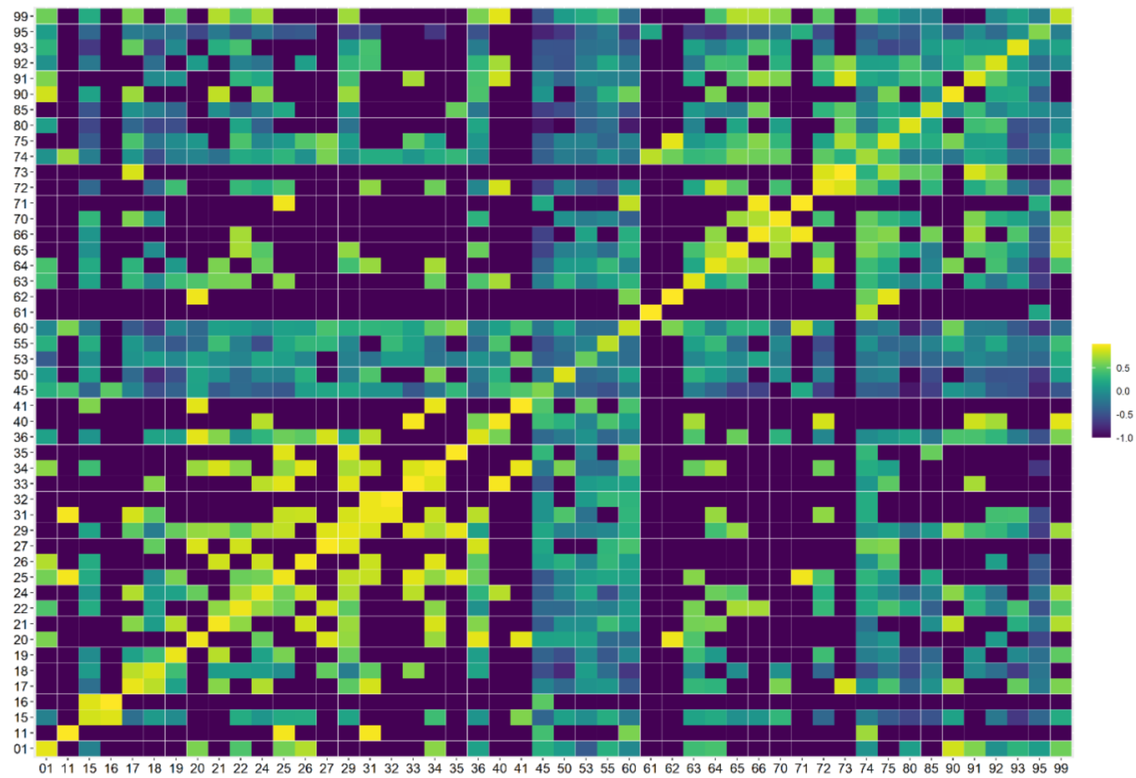

Figure 6. Heat map of industries of the Sociodemographic Economic Industry's Classification (CAES in Spanish) based on data from the Permanent Household Survey (EPH)

We can express this measure of similarity between occupations through a weighted graph, as shown in Figure 7, based on all occupational codes, except for managers and high-ranking officials in the government, and their ties to other codes. The width of the lines between a pair of nodes is proportional to the size of the relationship between the nodes, while the size of the node represents the strength of each node –where the strength is the weighted graph's adaptation of the node degree, which shows the sum of all the links to other occupations, taking into account the intensity of the attraction.

In the figure, we can see that that some occupations are closely linked to many other occupations. This is the case, for example, of consulting, recreational services, administrative management, planning and management control, and education, among others. By contrast, two occupations that explained 14% of the total employment in 2019 (construction and domestic services) show little connectivity to other nodes. In the case of domestic services, its only important links are to human care and assistance and non-domestic cleaning services, while Construction is linked to the storage of supplies and various social services.

Figure 7. Weighted graph of occupation codes the transitions from the Permanent Household Survey (EPH). Negative similarity values are excluded. Maximum value of the relationship between both occupation nodes are shown. The width of the link between the nodes represents the value of the similarity measure. Showing occupations with at least 50 worker transitions detected.

Although self-loops (self-similarities) are excluded from the graph, it is interesting to observe that transitions between same occupations are the rule rather than the exception. As an example, 8 of the

10 edges with the highest similarity value and almost half of all job changes occur between the same occupations.

What can we say about central occupations in the constructed graph? The distribution of strength can give us information about what occupations are more directly related to other occupations (Figure 8): the occupations with greater *strength* are those linked to administrative management, planning and management control. In addition to the aforementioned domestic services and construction, transport and industrial and craft productions also have low "attraction" to other occupational codes.

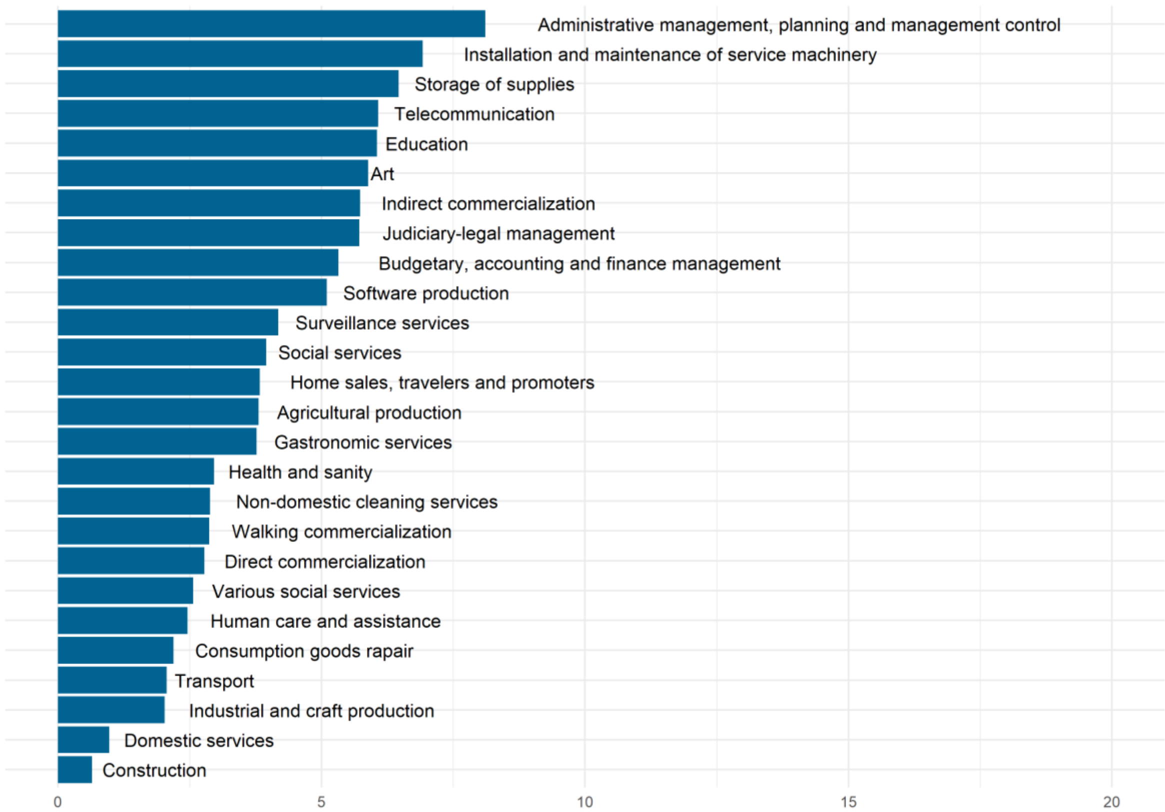

Figure 8. Strength of each of the vertices (occupations) of the weighted graph of occupations based on the transitions detected in the Permanent Household Survey (EPH). Showing occupations with at least 50 worker transitions detected.

Figure 9 and 10 show the same information for industry transitions. The main difference is that manufacturing is subdivided between several different industry codes that are heavily interconnected in our similarity measure. Most of the codes in the center of the graph (Figure 9) are indeed manufacturing codes, such as Leather Products, Textile fabrics, Chemical Products or Furniture. One key takeaway from both industry and occupation transition analysis is that construction and domestic services are mostly isolated in the short-term job-to-job transitions, both analyzing them as occupations or industries.

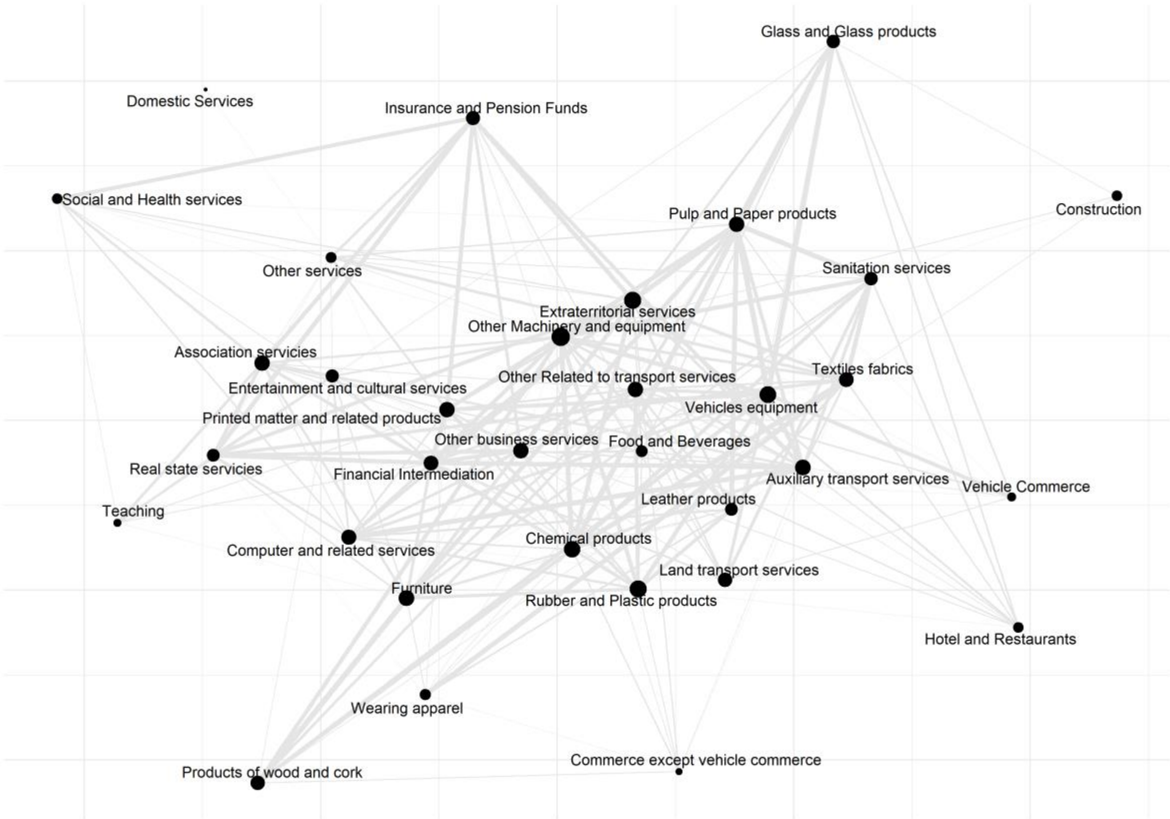

Figure 9. Weighted graph of industry codes the transitions from the Permanent Household Survey (EPH). Negative similarity values are excluded. Maximum value of the relationship between both industry nodes are shown. The width of the link between the nodes represents the value of the similarity measure. Showing industries with at least 50 worker transitions detected.

The analysis of the strength nodes at the industry level has the same bias than when we analyzed the graph. Manufacturing codes dominate the main positions of the strength ranking (Figure 10) mainly because they are heavily interconnected among each other. Some transport sectors, financial intermediation and real estate services stand out, besides the manufacturing codes. Domestic services, commerce, education, and construction show low connection to other industry codes.

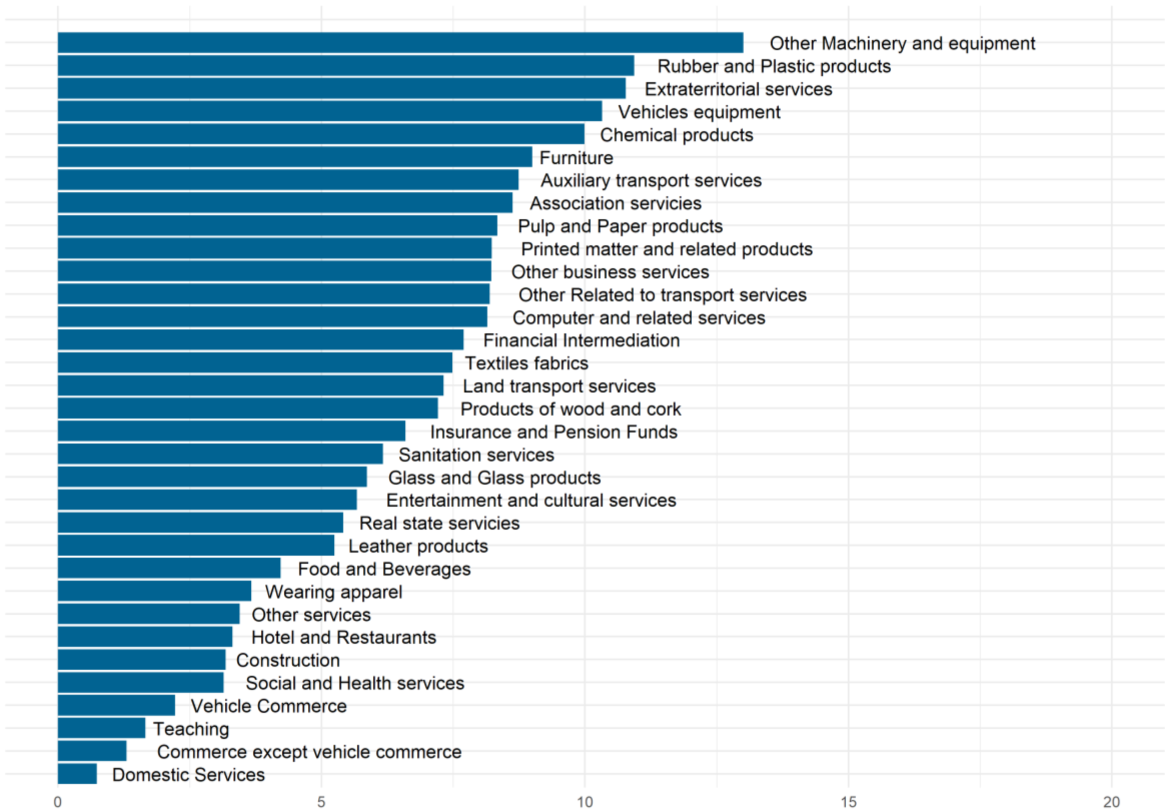

Figure 10. Strength of each of the vertices (industries) of the weighted graph of industries based on the transitions detected in the Permanent Household Survey (EPH). Showing industries with at least 50 worker transitions detected.

The specificity of human capital

Several studies have examined in depth how general our human capital is (Poleteav and Robinson, 2008; Kambourov and Manovskii, 2009; Gathmann & Schonberg, 2010; Sulivan, 2010; Yamaguchi, 2012; Nedelkoska et al; 2018). By human capital, we mean a set of intangible knowledge and experience that makes it possible for us to perform tasks and occupations in a successful way.

Knowing how specific our human capital is becoming especially relevant from the viewpoint of labor reconversion and reskilling. If human capital was not specific, it would be easy for a worker to find a new career path in the event demand for her industry, occupation or tasks began to decline. But, if moving between industries, occupations or tasks comes at a significant cost in terms of worker productivity, then the costs of reconversion are high and a set of informed policies that allow for a smoother transition between jobs through adequate reskilling provides an invaluable tool, particularly when a crisis hits sectors deeply and unevenly, and labor turnover peaks.

Ideally, testing the link about turnover and human capital loss requires some measure of worker productivity before and after a move, correcting for the seniority she has in each of the industry sector or occupation. Unfortunately, productivity is not usually captured in household surveys or administrative data, and even if they do, the purely observational nature of these measures would fall short of what we need for a rigorous test.

Because of that, empirical research often resorts to an imperfect measure of labor productivity: the wages perceived by the worker. Along those lines, Kambourov and Manovskii (2009) find that, after time-monitoring the employer and the activity sector, seniority in the same occupation explains a significant part of the wage received. Gathmann and Schonberg (2010) go further and, based on administrative data from Germany with a disaggregation of tasks at the individual level, manage to identify a specific effect associated with the time that a worker has been performing the same task. Closer in time, Nedelkoska et al (2018) find that similarity between the occupations, measured in terms of the task-intensity, is a good predictor of wage changes.

The Permanent Household Survey allows us to perform a similar exercise for a developing economy: Argentina. We build on the strategy used in Gathmann and Schonberg (2010) and Nedelkoska et al (2018). The key concept is that the wage at the new job should be more correlated to the last job wage when human capital is more similar between the two jobs. We measure human capital similarity between the two jobs through our measure of industry and occupational similarity based on the job-to-job transitions and estimate different variations of the following model:

$$\ln\left(HRW_{p,t,j}\right) = c + \beta_1 * \ln\left(HRW_{p,t,i}\right) + \beta_2 * Rij + \beta_3 * Rij * \ln\left(LHRW_{p,t,i}\right) + X'_{p,t} + Y'_{p,t,j} + Z'_{p,t,j} \; + T_t$$

Where $HRW_{p,t,j}$ is the hourly real wage of person p at time t in occupation or industry j (after the transition), $HRW_{p,t,i}$ is the hourly real wage of person p at time t in occupation or industry i (before the transition), Rij is the similarity measure between occupation/industry i and j, X'_{pt} is a set of characteristic of the worker (such as education and age), $Y'_{p,t,i}$ is a set of variables related to the job last position held (such as hours worked and labor relationship), $Z'_{p,t,j}$ is a set of variables related to the job new job position (such as hours worked and labor relationship), and T_t are time dummies[5].

Table 2 shows the estimation of the main parameters in three regressions: 1) only includes occupational similarity, 2) only includes industry similarity, while 3) includes both similarity measures. The relevant parameters, the one in which we are interested in, is the interaction between the lagged hourly real wage and the similarity measure between occupations (RijOccupations) and between industries (RijIndustries).

[5] In the Annex 3 we show the full specification of the three regressions.

	Occupational similarity (1)	Industry similarity (2)	Both similarity measures (3)
$\ln(LHRW_{p,t,j})$	0.20 ***	0.23 ***	0.20 ***
	(0.01)	(0.01)	(0.02)
RijOccupations	-0.87 ***		-0.49 *
	(0.14)		(0.19)
RijIndustries		-0.80 ***	-0.56 ***
		(0.12)	(0.16)
$\ln(LHRW_{p,t,j})$* RijOccupations	0.16 ***		0.10 **
	(0.02)		(0.03)
$\ln(LHRW_{p,t,j})$* RijIndustries		0.14 ***	0.09 ***
		(0.02)	(0.03)
Intercept	4.65 ***	4.43 ***	4.61 ***
	(0.22)	(0.24)	(0.24)
Nobs	5976	5575	5382
r.squared	0.33	0.33	0.33
adj.r.squared	0.32	0.33	0.33

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table 2. Output of three regressions that have as a dependent variable the natural logarithm of the real hourly wage in the new occupation. (1) Uses the occupational similarity only, (2) uses industry similarity only, (3) combines both similarity measures. Annex 3 shows the full specifications and results for these three equations.

In all three regressions we find positive and significant coefficients for the similarity of the occupations and industries, both in their individual regressions and in the one that includes both. We interpret this result as evidence of specificity of occupational and industry human capital specificity: the more similar that the occupations or industries are, the more the lagged and the new real wages are positively correlated.

Conclusion

The ability of workers to reconvert to new labor demands at the lowest possible cost has long been at the core of the political debate on the response to globalization, to technological change and, more recently, to the devastating effects of the COVID-19 crisis. A key question on this front in how general or "portable" our human capital is and, given the lack of total transferability, how we should inform our reskilling efforts based on the transitioning worker´s human capital.

In this paper, we made a contribution by assessing the specificity of human capital in a developing country. First, we created a similarity measure of industries and occupations based on the actual flows of workers' transitions. Then, we extracted two relevant stylized facts: 1) older people and more educated workers tend to transition among occupations and industries that are deemed as similar based on our measure, 2) around half of the transitions are within the same occupation or industry, showing that there is a revealed preference by the workers to stay in the same industry or occupation.

We also analyzed the weighted graphs resulting from our similarity measures. We found that key industries and occupations in developing countries, such as construction and domestic services, are not heavily connected to other occupations or industries. They appear in the last positions of the occupations and industries when ordered by their strength in the weighted network.

Finally, using our similarity measure in a regression explaining the hourly real wage at the new job position, we found that the more similar the initial occupation (industry) is to the new occupation (industry), the closer is the new wage to the previous one. We interpret these stylized facts and econometric results as indication of the specificity of human capital, both at the occupational and the industry level.

This study has obvious limitations that need to be considered. First, these are short-term transitions: it would be useful to calibrate the Occupation Space based on a long data panel to follow longer job trajectories. Also, this is observational evidence that needs to be interpreted carefully. That said, we believe the methodology and results prove that this line of research is a valuable input in the design of active labor market policies.

Annex 1 – Occupational and industrial codes

Code	Description
0	Executive officials
1	Legislative officials
2	Judiciary officials
3	Agencies, firms and state managers
4	Social institution managers
5	Small firm managers
6	Medium firm managers
7	Big firm managers
10	Administrative management, planning and control management
11	Judiciary-legal management
20	Budgetary, accounting and finance management
30	Direct commercialization
31	Home sales, travelers and promoters
32	Indirect commercialization
33	Door-to-door commercialization occupations
34	Transport occupations
35	Telecommunication
36	Storage of supplies
40	Health and sanity
41	Education
42	Research
43	Consulting
44	Claims prevention

45	Mass communication
46	Social services
47	Surveillance services
48	Police services
49	Armed forces
50	Art
51	Sport
52	Recreational services
53	Gastronomic services
54	Accomodation and tourism services
55	Domestic services
56	Non-domestic cleaning services
57	Human care and assistance
58	Various social services
60	Agricultural production
61	Livestock production
62	Forestal production
63	Bee and poultry production
64	Fishing production
65	Hunting
70	Extractive production
71	Energy, gas and water production
72	Construction
80	Industrial and craft production
81	Software production
82	Consumption goods rapair

90	Machinery instalation and maintenance
91	Productive technological development
92	Installation and maintenance of service machinery

Table A1. Occupations of the National Occupational Code (CNO) 2001.

Code	Description
1	Agriculture/Livestock
2	Wood
5	Fishing
10	Coal and Lignite
11	Crude petroleum and Natural Gas
12	Uranium and Thorium Ores
13	Metal ores
14	Others OIL&GAS
15	Food and Beverages
16	Tobacco Products
17	Textiles fabrics
18	Wearing apparel
19	Leather products
20	Products of wood and cork
21	Pulp and Paper products
22	Printed matter and related products
23	Refined petroleum products
24	Chemical products
25	Rubber and Plastic products
26	Glass and Glass products

27	Basic Metals
28	Fabricated metal products
29	Other Machinery and equipment
30	Office, Accounting and Computing machinery
31	Electrical machinery and apparatus
32	Radio and TV equipment
33	Medical Appliances, precision and optical instruments
34	Vehicles equipment
35	Other transport equipment
36	Furniture
37	Recycling
40	Gas and Electricity
41	Water
45	Construction
50	Vehicle Commerce
53	Commerce except vehicle commerce
55	Hotel and Restaurants
60	Land transport services
61	Water transport services
62	Air transport services
63	Other Related to transport services
64	Auxiliary transport services
65	Financial Intermediation
66	Insurance and Pension Funds
67	Auxiliary services to financial activity
70	Real state servicies

71	Transport equipment rental
72	Computer and related services
73	Research and development services
74	Other business services
75	Public administration
80	Education
85	Social and Health services
90	Sanitation services
91	Association servicies
92	Entertainment and cultural services
93	Other services
95	Domestic Services
99	Extraterritorial services
0	Other activities

Table A2. Sociodemographic Economic Industry’s Classification (CAES Mercosur)

Annex 2 – The *seniority rule* method

The Permanent Household Survey (EPH) of Argentina has different variables related to seniority, depending on the worker´s contract and activity sector. If the worker is a salaried but not employed in domestic services, then question PP07A asks the following:

How long have you been employed in this job continuously? (PP07A)

1 = less than a month
2 = 1 to 3 months
3 = 3 to 6 months
4 = 6 to 12 months
5 = 1 to 5 years
6 = more than 5 years

If the worker is independent, but not employed in domestic services, the PP05H asks the following question:

How long have you been employed in this job continuously? (PP05H)

1 = less than a month
2 = 1 to 3 months
3 = 3 to 6 months
4 = 6 to 12 months
5 = 1 to 5 years
6 = more than 5 years

Finally, if the worker is employed in domestic services, the PP04B3_MES and PP04B3_ANO give us information on the years and months since the worker have been working in the main house in which he or she is currently working.

For each worker we calculate the seniority on the job considering these three scenarios. We always assume that the experience is equal to the maximum value in each bracket to minimize the risk of detecting too many job switches. For each person, we identify a change in employment whenever the time since he or she entered the sample is greater than the seniority variable we created.

Annex 3 – Full regression specification

	Occupational similarity	Industry similarity	Both similarity measures
Intercept	4.65 ***	4.43 ***	4.61 ***
	(0.22)	(0.24)	(0.24)
$\ln(LHRW_{p,t,j})$	0.20 ***	0.23 ***	0.20 ***
	(0.01)	(0.01)	(0.02)
RijOccupations	-0.87 ***		-0.49 *
$\ln(LHRW_{p,t,j})$* RijOccupations	0.16 ***		0.10 **
	(0.02)		(0.03)
RijIndustry		-0.80 ***	-0.56 ***
		(0.12)	(0.16)
$\ln(LHRW_{p,t,j})$* RijIndustry		0.14 ***	0.09 ***
		(0.02)	(0.03)
	(0.14)		(0.19)
Occupational category of lagged employment Case base: informal waged worker			
Formal waged worker	0.01	0.00	0.01
	(0.02)	(0.03)	(0.03)
Independent worker	0.05 *	0.05 *	0.04
	(0.02)	(0.02)	(0.02)
Occupational category of present employment Case base: informal waged worker			
Formal waged worker	0.37 ***	0.38 ***	0.37 ***
	(0.02)	(0.03)	(0.03)
Independent worker	0.04	0.05 *	0.03
	(0.02)	(0.02)	(0.02)
Time dummies (year) Case base: 2003			
2004	0.01	0.05	0.04
	(0.21)	(0.22)	(0.22)
2005	0.06	0.08	0.08

	(0.21)	(0.22)	(0.22)
2006	0.17	0.20	0.20
	(0.21)	(0.22)	(0.22)
2007	0.17	0.21	0.20
	(0.21)	(0.22)	(0.22)
2008	0.09	0.12	0.11
	(0.21)	(0.22)	(0.22)
2009	0.21	0.26	0.25
	(0.21)	(0.22)	(0.22)
2010	0.21	0.24	0.23
	(0.21)	(0.22)	(0.22)
2011	0.25	0.26	0.26
	(0.21)	(0.22)	(0.22)
2012	0.30	0.33	0.33
	(0.21)	(0.22)	(0.22)
2013	0.30	0.32	0.32
	(0.21)	(0.22)	(0.22)
2014	0.08	0.10	0.10
	(0.21)	(0.22)	(0.22)
2015	0.25	0.26	0.27
	(0.21)	(0.22)	(0.22)
2016	-0.06	-0.04	-0.03
	(0.23)	(0.25)	(0.24)
2017	0.25	0.25	0.27
	(0.21)	(0.22)	(0.22)
2018	0.10	0.10	0.11
	(0.21)	(0.22)	(0.22)
2019	-0.11	-0.08	-0.07
	(0.21)	(0.23)	(0.22)
AGE (years)	0.00 ***	0.00 ***	0.00 ***
	(0.00)	(0.00)	(0.00)

Maximum educational level			
Case base: Complete elementary school			
High school dropout	0.06 **	0.07 **	0.06 **
	(0.02)	(0.02)	(0.02)
Complete High school	0.14 ***	0.15 ***	0.14 ***
	(0.02)	(0.03)	(0.03)
Some university studies	0.23 ***	0.25 ***	0.23 ***
	(0.03)	(0.03)	(0.03)
Complete university studies	0.44 ***	0.50 ***	0.45 ***
	(0.03)	(0.04)	(0.04)
Weekly hours of employment (present employment)	-0.02 ***	-0.02 ***	-0.02 ***
	(0.00)	(0.00)	(0.00)
Weekly hours of employment (lagged employment)	0.01 ***	0.01 ***	0.01 ***
	(0.00)	(0.00)	(0.00)
nobs	5976	5575	5382
r.squared	0.33	0.33	0.33
adj.r.squared	0.32	0.33	0.33

*** p < 0.001; ** p < 0.01; * p < 0.05.

References

Balassa, B. (1986). Intra-industry specialization: a cross-country analysis. *European Economic Review*, *30*(1), 27-42.

De Raco, S. A., & Semeshenko, V. (2019). Labor mobility and industrial space in Argentina. *Journal of Dynamics & Games*. *6*(2), 107-118.

Gathmann, C., & Schönberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, *28*(1), 1-49.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the national academy of sciences*, *106*(26), 10570-10575.

Hidalgo, C. A., Klinger, B., Barabási, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482-487.

Kambourov, G., & Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, *50*(1), 63-115.

Montané, M., & Sartorio, L. (2018). Tecnología y generación de empleo en el siglo XXI: desafíos y propuestas. Premio de la Academia Nacional de Ciencias Morales y Políticas, 2019.

Nedelkoska, L., Diodato, D., & Neffke, F. (2018). *Is Our Human Capital General Enough to Withstand the Current Wave of Technological Change?* (93a). Center for International Development at Harvard University.

Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, *34*(3), 297-316.

Neffke, F. M., Otto, A., & Weyh, A. (2017). Inter-industry labor flows. *Journal of Economic Behavior & Organization*, *142*, 275-292.

Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. *Labour economics*, *17*(3), 567-580.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, *30*(1), 1-53.