Can the gendered sorting of occupations explain wage differentials across educational levels in the US?

Alícia Adsera (Princeton U), Federica Querin (University Bologna),Varun Satish (Princeton U)

(Extended Abstract. PLEASE see also preliminary draft attached)

Background and Motivation

The gender wage gap, which once appeared to be narrowing rapidly, has hit a plateau in recent years. The persistence of gender earnings gaps, despite women having higher levels of education, runs in contrast to predictions from models of human capital accumulation. To address this, we ask whether men and women occupy different types of jobs even within the same educational level.

Education Differentials: Transformations of the Labor Market

Important structural changes in the labor market starting in the 1980s affected workers and sectors differentially. For example, workers with higher or lower educational attainment either benefited or hurt from skill biased technological change and mechanization of jobs. To understand those differences, David Autor et al (2003) distinguished among jobs that involved different degrees of either cognitive or non-cognitive routine tasks or non-routine ones (differentiating further within them among analytical and interpersonal tasks). Cognitive routine jobs for example, which have been significant for women, have suffered due to technological changes and computerization

Educational Differentials in Occupational Sorting

Highly educated workers often find themselves in jobs that demand more personal relationships, possess unstructured work environments, and involve significant time pressure. These jobs tend to have a non-linear hourly wage structure, rewarding long hours. However, women may avoid such occupations or experience substantial wage penalties for reducing their effort. In such scenarios, the gender wage gap widens. The COVID-19 pandemic thrust work from home (WFH) into the spotlight. Occupations that do not require direct machinery use and entail limited face-to-face interaction are more amenable to remote work. On the flip side, less flexible jobs that necessitate machinery operation are more common among low-skilled workers. The extent to which WFH

characteristics reduce stress, enable better work-family balance, or simply reduce commuting costs remains unclear.

Educational Differentials and Occupational Characteristics

In addition to sectoral differences, occupational characteristics can contribute differentially to the gender wage gap across education levels. A key attribute explored in our work is "contact with others". The focus has primarily been on occupations requiring face-to-face contact, but advancements in information technology have extended the realm of interpersonal working relationships. In contrast, leadership roles, typically dominated by men, yield higher wage returns. We also look at the degree of autonomy in decision-making in the occupation, usally associated with managerial positions, IT is distinct from leadership, as it doesn't necessarily entail extensive contact with others. While less-structured occupations may demand more decision-making and offer more freedom, they also come with greater responsibilities.We hypothesize that these responsibilities may lead to higher wage returns, particularly for highly educated workers. In contrast, some jobs are considered highly inflexible due to the need for direct machinery use and physical presence. These machine-dependent occupations, often male-dominated, do not necessarily yield higher wages. Finally, time pressure and deadlines, which often conflict with family needs. The impact of time pressure on workers, especially mothers, varies based on individual circumstances, such as the sharing of household tasks. We estimate whether that occupational characteristic explains part of the wage differential.

Data and Methods

Data

This study employs data from the American Community Survey (ACS) 2011-2015 pooled sample, which provides information on occupation with 6-digit SOC codes which we can match directly to the O*NET dataset, which provides granular descriptions of job characteristics. We study men and women who are 16-65 years old, not self-employed and fairly attached to the labor market (defined as working at least 20 hours a week and 26 weeks a year), We group respondents by mutually exclusive educational groups:

Occupational Characteristics

For each currently employed individual, we match occupational characteristics corresponding with the 6-digit SOC with information retrieved from the O*NET database. The database, available online https://www.onetonline.org/, has detailed descriptions of the requirements and work content and characteristics of over 900 occupations. We use O*NET release 25.0 (August 2020to construct the O*NET factors that are normalized to have mean of zero and a standard deviation of one at the occupation-level, with weights based on the share of the ACS analytical sample working in a given occupation. We use five occupational indicators for work context provided by O*NET and briefly described here:

1. *Contact* with others captures the importance of communicating with others (be it colleagues, external customers, or the public) without a clear hierarchical structure in the relationship. It reflects and extends factors presented in Goldin (2014) who suggests more contact means less flexibility as others need the worker to be present to carry out her job.
2. Second, our measure of *leadership* also reflects strong interpersonal contact with others. However, unlike contact with others, items in this factor explicitly refer to relationships that are hierarchical in nature and where the worker has a coordinating and guiding role (e.g., coordinating the work and activities of others, guiding, directing, and motivating subordinates, etc.). Therefore, there is still a need for interpersonal relationships to fulfil one's job, but the worker has more latitude in directing them.
3. Third, *autonomy* relates to the importance, frequency, and freedom to take up the role of decision-maker. While this measure could be correlated with leadership, it is distinct from it by the absence of an explicit relationship with others.
4. Fourth, *machine-dependency* is built by selecting from the measures of inflexibility suggested in the literature the items pertaining specifically to the operation of mechanized devices, moving objects, and the control of machines and processes (see for example Albanesi and Kim 2021). This factor explicitly excludes computers as machines and therefore captures features of blue-collar occupations more than the other factors.
5. Lastly, *time pressure* captures the time-sensitivity of the job, measured by how often the job requires the worker to meet strict deadlines, ranging from never to every day.

Methods: Our analyses include, separately by the educational groups:

1. Models on the prevalence, by gender, of each of the occupational characteristics in jobs held by respondents in the analytical sample.
2. Estimates of differences in log hourly wages by gender. We control for occupational characteristics in levels and interacted by gender to see whether the returns (in terms of wages) to the prevalence of those characteristics varies by gender.
3. Oaxaca-Binder decomposition to show how occupational characteristics contribute to the gender wage gap.

Preliminary Results

In the attached draft of the paper, we include some preliminary results that confirm gendered prevalence in these occupational characteristics, with women performing jobs with higher contact with others and less machine use. Table 1 below reports the female coefficients in linear regressions, estimated separately by education, where the dependent variable is each of the five occupational skills.

- Contact with others is the only occupational characteristic that has more prevalence amongst women than men (positive coefficients throughout).
- Women are especially less prevalent in occupations characterized using machines, with the gap with men being smaller for the lowest education group (where both men and women are more likely to hold these jobs) and the highest education group (where both men and women are unlikely to hold these jobs).
- Except for leadership for those who did not complete high school and time pressure for those who completed college, the other female coefficients have the same sign across all education groups. This suggests that, albeit to different extents, women are consistently over- or underrepresented in certain occupational characteristics regardless of education.
- The completed college group has among the smallest female coefficients in absolute terms, suggesting a smaller gender prevalence gap (i.e. less occupational sorting).

Second, we document educational gradients in the wage returns to these characteristics.

Third, results from Oaxaca-Binder decompositions highlight how our models explain gender differentials more for highly educated workers than for those with less education, underscoring the importance of further research specifically on workers who do not have a college degree.

Additional Work

Considering the preliminary results, we are interested in exploring further educational differences and look for specific occupational characteristics that are more relevant for those who are lower educated. In that regard we are interested in using data-driven approaches such as clustering techniques to identify "groups" of occupational characteristics (using a larger set of what is available through O*NET) that seem to matter for predicting earnings amongst the lower-educated. Our hope here is that we will be able to uncover some "unobserved" groups of characteristics that explain wage differentials. The final version of the paper will include data from previous periods to track how the gendered sorting of occupational characteristics has evolved over time.

Table 1: Difference Between Men and Women in the Prevalence of Occupational Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample	Less than HS	HS, no college	1-2 years college	4 years college	5+ years college
Contact	0.24***	0.19***	0.36***	0.29***	0.13***	0.05***
	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Autonomy	-0.15***	-0.22***	-0.15***	-0.15***	-0.08***	-0.08***
	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Time pressure	-0.11***	-0.25***	-0.23***	-0.12***	0.01***	0.02***
	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Leadership	-0.09***	0.03***	-0.06***	-0.11***	-0.07***	-0.04***
	(0.001)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)
Machine	-0.30***	-0.27***	-0.49***	-0.40***	-0.13***	-0.04***
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
N	5,569,817	311,801	1,787,116	1,393,108	1,249,317	829,206

(female coefficient for prevalence models controlling for other characteristics)