First in family university graduates on the labor market: the role of selection to firms and occupations

Anna Adamecz[1], Morag Henderson[2] and Nikki Shure[3]

28 February 2023

Work in progress

Abstract

Recent studies have revealed that "first in family" (FiF) graduates suffer a wage penalty in the labor market compared to graduates whose parents are graduates, and this penalty is larger for women. Using linked employer-employee administrative panel data from Hungary, this paper sheds light on two potential channels behind this phenomenon: the role of selection to firms and occupations. We show that in early career, FiF graduate women earn 4.1% less per hour than graduate women whose parents are graduates, conditional on pre-university educational attainment, university course, degree type, industry, occupation, firm size, and local labor market fixed effects. The male FiF wage gap is 1.1%. FiF graduates of both genders tend to work at firms that pay lower wages on average, produce lower 'value added' per employee, and have lower firm-specific wage premia than non-FiF graduates. Interestingly, within-occupation-at-the-same-firm, the FiF wage gap is similar for men and women (1.8-1.9%). Investigating selection to occupations, we link the cognitive ability level requirements of occupations from the O*NET database and find that FiF graduate women tend to work at lower-cognitive-ability-level jobs than non-FiF graduate women, which is not true for men. We conclude that selection to firms and occupations are important factors behind the FiF wage gap, and they are especially crucial for women. Taking a step further we also investigate how female-friendly the firms are where FiF and non-FiF graduates work by looking at what happens to their female employees after maternity leave. We find that although FiF women are less likely to have a child at age 25 than non-FiF graduate women, the firms where they work are more female-friendly than the firms where non-FiF graduates work. Still, controlling for these firm characteristics does not explain a meaningful share of the FiF female wage gap. Hence, we do not find evidence that FiF graduate women earn less because they would trade-off financial rewards for maternity-related flexibility.

[1] Corresponding author. Institute of Economics, Centre for Economic and Regional Studies, Toth Kalman u. 4, 1097 Budapest and UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA. ORCID: 0000-0001-6408-5830. E-mail: a.adamecz-volgyi@ucl.ac.uk

[2] UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA. ORCID: 0000-0003-0107-4899. E-mail: morag.henderson@ucl.ac.uk

[3] UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA and Institute of Labor Economics (IZA), Schaumburg-Lippe-Str. 5-9, D-53113 Bonn. ORCID: 0000-0002-1270-4131. E-mail: nikki.shure@ucl.ac.uk

Keywords: first in family graduates, linked employer-employee data, wage gaps, intergenerational mobility, gender economics, O*NET database

JEL-codes: I24, I26, J16, J24, J31, J71

Acknowledgements: Anna Adamecz is grateful for the support of OTKA FK-138015. The authors thank Ágota Scharle, Balázs Reizer, Virág Illyés, Rita Pető, and the participants of seminar and conference presentations at the KRTK KTI for useful comments.

1. Introduction

Higher education (HE) brings substantial benefits to graduates and society: university graduation increases the probability of employment, brings higher wages, and contributes to positive health outcomes (Oreopoulos and Petronijevic 2013). Increasingly, technological complexity and automation have also increased the demand for certain kinds of skills, which may make going to university an even better investment. Recent predictions expect that half of all new jobs in the EU between 2013 and 2025 will require an HE qualification, and that employment will decrease in occupations that require medium- or low-level qualifications (European Commission 2017). Driven by these phenomena, the share of university graduates in the EU has doubled in the last 25 years. The expansion of HE implies that a fair share of recent university graduates is the “first in their families” or first-generation to go to university (FiF). In England, for example, about two-thirds of recent university graduates are FiF (Henderson, Shure, and Adamecz-Völgyi 2020).

Despite this high share of FiF graduates, there is limited evidence on how they fare on the labor market. In the US, Nunez and Cuccaro-Alamin (1998) find no wage difference between first-generation and second-generation graduates one year after graduation in the '90s. In this same period, Thomas and Zhang (2005) find a small FiF penalty shortly after graduation, increasing to about 4% four years post-graduation. Also in the US, Manzoni and Streib (2019) find a 10% FiF wage gap 10 years after graduation that decreases to 3-4% after controlling for race, fertility, early educational attainment and labor market choices (industry, occupation, hours worked, and location). Using data from England, Adamecz-Völgyi, Henderson, and Shure (2022) find that young female FiF graduates suffer a wage penalty of 7.4% at age 25 compared to female graduates whose parents are graduates, while men do not.

This paper further investigates the FiF wage gap using linked employer-employee administrative panel data from Hungary. Our data, which covers 50% of the Hungarian population, allows us to link graduates to the firm where they work. We also observe the financial statements of these firms, including data on their sales revenues, ‘value added’, as well as the individual employment and healthcare data of the firms’ employees. Thus, we can compare the characteristics of firms where FiF and non-FiF graduates work, which has not been previously done in the literature. This is important since there are large differences in earnings within occupations between firms and selection into firms may be related to social background.

We use data on two school cohorts born in 1991-1993, whose employment outcomes are observed until age 25-26. We estimate Mincer-type wage models to look at the FiF gap in log hourly wages, while we control for pre-university educational attainment (test scores), industry, occupation, firm size, and local labor market fixed effects. We find that FiF graduate women suffer a 4.1% conditional wage penalty compared to female graduates whose parents are graduates, while the FiF gap among men is 1.1%. As we observe all employment spells of these young people every month between 2003 and 2017, we also show that the FiF wage gap emerges right after graduation.

We show that FiF graduates tend to work at "worse" firms in terms of average wages, value added per employee, sales revenues per employee, and firm-specific wage premia (that is conditional on the distribution of workers). In most cases, these FiF gaps in the performance indicators of firms are larger for women. The performance indicator of firms, especially firm-level per capita value added and firm-specific wage premia, explain 55-83% of the FiF wage gaps. Interestingly, within-occupation-at-the-same-firm, the FiF wage gap is similar for men and women (1.8-1.9%).

To try to understand the role of selection to occupations, we use the O*NET database to link the cognitive ability requirement of jobs to occupations on a 4-digit level. We find that among women, FiF graduates tend to work at lower-cognitive-ability jobs than non-FiF graduates, while we do not see a difference between the cognitive ability requirements of jobs where FiF and non-FiF male graduates work.

Lastly, we ask the question why FiF graduates work at "worse" firms. One potential answer could be that they value something else more than financial rewards, and thus this is their "choice". We test this hypothesis by looking at the role of children, and the role of child-related policies at the firms in the FiF wage gap. We link individual-level healthcare and income transfer data to all women in the sample working at all firms. This allows us to identify if (and when) women give birth, and what happens after their maternity leave: whether they go back to the same firm (and occupation), how is their relative wage compared to their wage before maternity leave, and how is the share of women and mothers at the firm where they work. Using this information, we construct firm-level measures to proxy how "female-friendly" a firm is. We find that although firms where FiF women work tend to be more female-friendly, controlling for these measures does not change the female FiF penalty in a meaningful way.

We make four contributions to the literature. First, to the best of our knowledge, we are the first to look at the labor market outcomes of FiF university graduates in a non-Anglo-Saxon country. In Hungary, or in any other Central-Eastern-European countries, intergenerational educational mobility is different from the US or in England. It might be that lower educational mobility (which we document) implies different selection mechanisms and leads to different results. Second, as opposed to the current literature that relies on self-reported earnings data, we observe wages in administrative data. Since we observe 50% of the population, we can exclude a potential bias coming from FiF graduates being more likely to over- or underreport their wages (or coming from measurement error in general). Third, we can link graduates to the firms where they work, where we also see 50% of all other employees working at the same firms, including their employment and (for women) fertility history during the observation period. Thus, we can investigate the role of selection to firms in the gendered FiF wage gap. Fourth, we are the first to investigate whether FiF graduates select to jobs that require lower levels of cognitive skills. This is important because sheds light on the phenomenon that FiF graduates undermatch in the labor market.

The rest of the paper is structured as follows. Section 2 introduces the data and the variables we construct. Section 3 presents descriptive statistics, Section 4 details our methods. We show our results in Section 5. Section 6 investigates the role of selection to firms. Section 7 concludes with a discussion.

2. Data

We use linked employer-employee administrative panel data (Admin3; Sebők 2019) covering a random sample of 50% of the Hungarian population born before 1 January 2003. This equates to about 45-50,000 people per birth year. For those born after June 1991, the data are also linked to the national school census, as well as information on centrally organized national exam scores, parental background, and information on higher education participation and earned degrees. Thus, we use the first two school cohorts already covered by the schooling data: those born between June 1991 and May 1993. These are young people who enrolled in elementary school either in September 1998 (those born between June 1991 and May 1992) or 1999 (those born between June 1992 and May 1993). We observe and/or construct the following variables.

Demographic variables: gender**,** year and month of birth, subregion (*járás*) as a proxy to capture the local labor market.

Higher education outcomes: Highest educational attainment, year of earning a degree; type of a university degree (BA, MA); course that we aggregated to STEMLEM/OSSAH/OTHER categories.

Parental education/being FiF: For those who took low-stake national examinations in grade 6, 8 or 10 (depends on the cohort and the absent rate, about 80% of students take them), we observe parental education from take-home surveys. We consider a young person FiF is neither of their parents earned a university degree (BA or above).

Grade 10 test scores: As a control variable, we use scores on national examinations in math and reading in grade 10 (age 16).

University aspirations: High school students were asked whether they wanted to continue their studies at a university at grade10.

Employment and wage data: We observe gross monthly wages, number of hours worked, occupation (2-3-4-digit ISCO occupation codes, 34-570 active categories), industry (letter code, 14 categories), and anonymized firm identifiers. We use log hourly wages as the main outcome variable. We do not correct nominal wages with the consumer price index, but we implicitly control for time effects by controlling for the cohort and the age of observation. The data are provided monthly between 2003 and 2017. At the end of the observation period, cohort members are aged 25-26for cohort 1 and 24-25 for cohort 2. If someone had multiple, parallel jobs, we look at the first (main) job (90% of people only have one job at a time).

Financial and other performance measures of firms: Out of the universe of all firms, double-accounting firms (roughly all firms with a sales revenue above 140,000 USD/year) are required to file income statements and balance sheet information annually for the tax authority. These reports are linked to our data using the same firm identifiers that are available for individual employment spells. We use yearly information on the number of employees, sales revenues, and construct log per capita value added (all revenues minus the costs of all intermediate inputs, which is the same as the sum of payments to labor and capital, plus taxes).

Firms-specific measures: Although our analytical sample is restricted to two school cohorts (the oldest of those having data on parental education and higher education outcomes), we observe 50% of all employees working at all firms. Thus, we can construct firm-specific measures to investigate the types of firms where FiF and non-FiF graduates work using the total universe of Hungarian firms. In particular, we construct

- average wage at the firm (excluding the wages of sample members so they would not affect the average);
- the share of women working at the firm;
- the share of women who have children at the firm (as a % of all women working at the firm);
- the average probability that if a woman working at the firm gives birth, she would come back after maternal leave to the same firm and occupation; and the relative log hourly wage after maternity leave compared to before maternity leave; and the average length of maternity leave at the firm. We use these measures as proxy variables for a female-friendly workplace.

Firm-specific wage premia: We follow the so called AKM-literature (see for example, Card, Cardoso, and Kline 2016) and estimate two-way fixed effects models with individual and firm fixed effects on the total sample of individuals and firms (i.e., the sample is not restricted to our two cohorts of interest). We normalize the estimated firm fixed effects and use them as a measure of firm-specific wage premia that is conditional on the individual productivity of the employees (or in other words, the distribution of employee characteristics of the firms). Appendix B provides more details about the estimation process and its underlying assumptions. We estimate general wage premia (that is constant within firms over employees and over time), and we also estimate gender-specific wage premia (within firms). We use these latter measures to construct the firm-level gender gap in the wage premia (male wage premia minus female wage premia) as a proxy for potential gender discrimination.

Having children: We define childbearing as either giving birth or getting child-related benefits. This information is only available for women. The data are linked individually to all inpatient and outpatient health care events between 2009-2017. Thus, we see if women gave birth in a public hospital. Furthermore, the data cover child-related income transfers that are usually available to the mothers of small children.

Level of required cognitive abilities in jobs: We merge data on the importance of cognitive abilities in occupations (at 4-digit ISCO code levels) from the O*NET database. We use the 20.1 edition[4] and combine all types of cognitive abilities that are required by occupations into one summary index. The index captures the average level of cognitive abilities that are important for each occupation. In O*NET, cognitive abilities are defined as "Abilities that influence the acquisition and application of knowledge in problem solving", and cover verbal abilities, quantitative abilities, problem solving, perceptual abilities, spatial abilities, and attentiveness to details. We take a simple average of all cognitive abilities (listed in Table B1 in Appendix B)[5].

The sample

Table 1 summarizes the coverage of the data by school cohorts. As mentioned above, these cohorts started elementary school in 1998 and1999, and they are covered in the data on average until age 25. We observe information on graduation for almost everybody in both cohorts (96-97%). For 87-91% of young people, we have grade 10 test score data, and for 64-71%, we have data on parental education. TBA: we need to look at selection to having data on parental education.

Table 1: The coverage of the Admin3 data by school cohort

School cohort born in	The earliest possible time to earn a university degree (3-year BA)	Employment data coverage until age on average (until 2017)	Share of individuals having data on			No. of individuals
			HE outcomes	Grade-10 test scores	Parental education	
June 1991 – May 1992	June 2013	25.6	0.96	0.87	0.64	
June 1992 – May 1993	June 2014	24.6	0.97	0.91	0.71	

Source: Admin3

[4] National Center for O*NET Development. *O*NET OnLine*. Retrieved February 10, 2023, from https://www.onetonline.org/
[5] Using the first predicted factor after a principal component analysis gives similar results as the simple mean.

3. Descriptive statistics

Intergenerational mobility and graduation

Table 2 summarizes some basic statistics. About 68-70% of young people completed a maturity exam. A maturity exam is an academic qualification taken at the end of high school (grade 12) at around age 18, and it is a prerequisite of HE application. About 36-38% spent at least one month at a university, and 17-19% graduated by age 25.

Table 2: HE statistics

School cohorts born in	Share of those who completed a maturity exam	Share of HE participants	Average HE participation length (months)	Share of graduates	Share of those still in HE in Dec 2017
June 1991 – May 1992	0.70	0.38	19.45	0.19	0.10
June 1992 – May 1993	0.68	0.36	17.66	0.17	0.14

Source: Admin3

As in England, about 70% of young people are potential FiF (Henderson, Shure, and Adamecz-Völgyi 2020), i.e. neither of their parent has a HE degree. Intergenerational mobility, however, is much lower in Hungary than in England. Out of potential FiF young people (those whose parents are not graduates), only 11-13% graduated, while the same share among children of graduate parents is 37-43% (Table 3). Consequently, the share of FiF among graduates is about two-thirds of that in England, at 41-42%.

Table 3: Potential FiF young people and FiF graduates

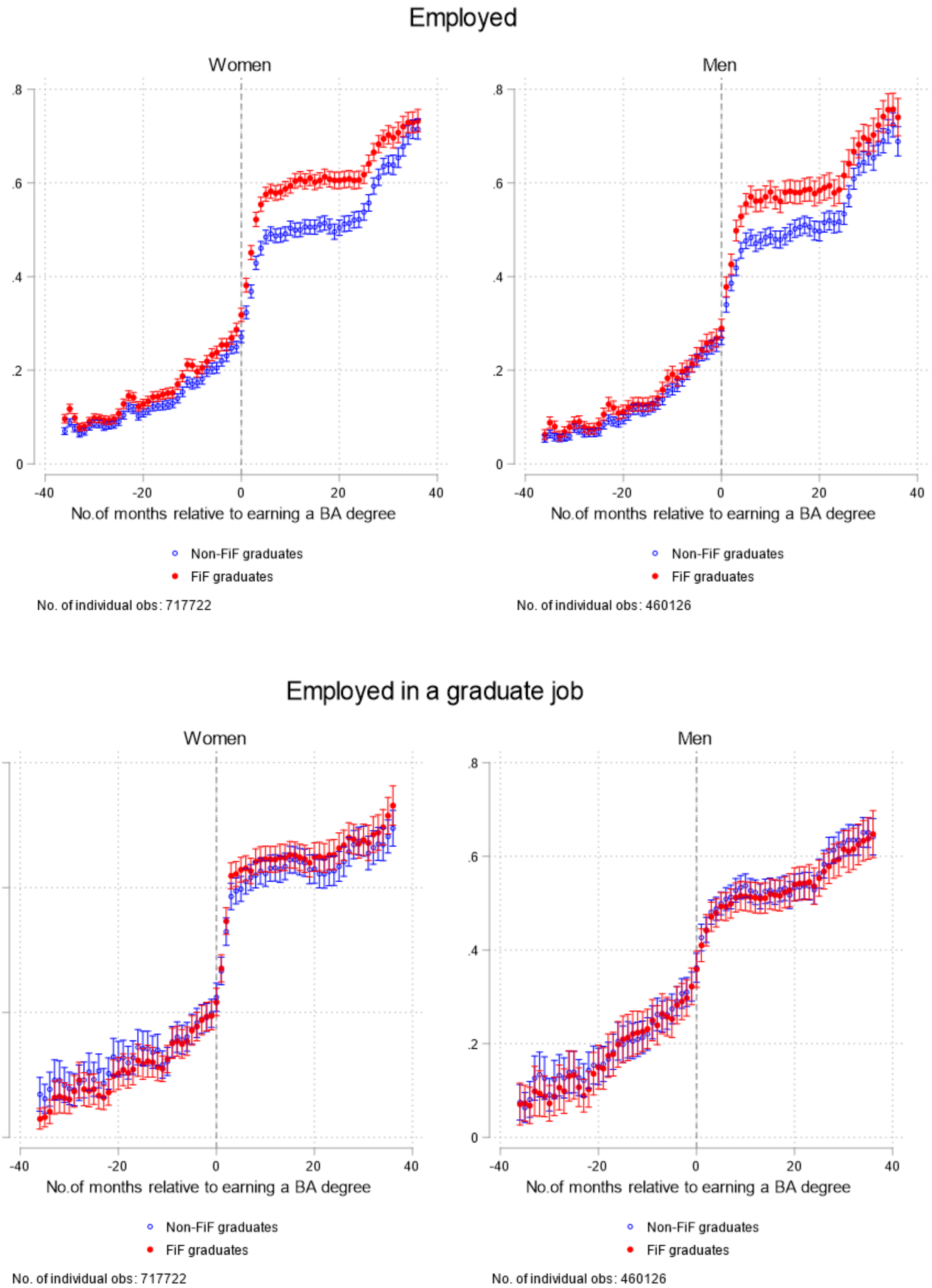
School cohorts born in	Share of potential FiF	Share of graduates among potential FiF	Share of graduates among children of graduate parents	Share of FiF among graduates
June 1991 – May 1992	0.71	0.13	0.43	0.42
June 1992 – May 1993	0.70	0.11	0.37	0.41

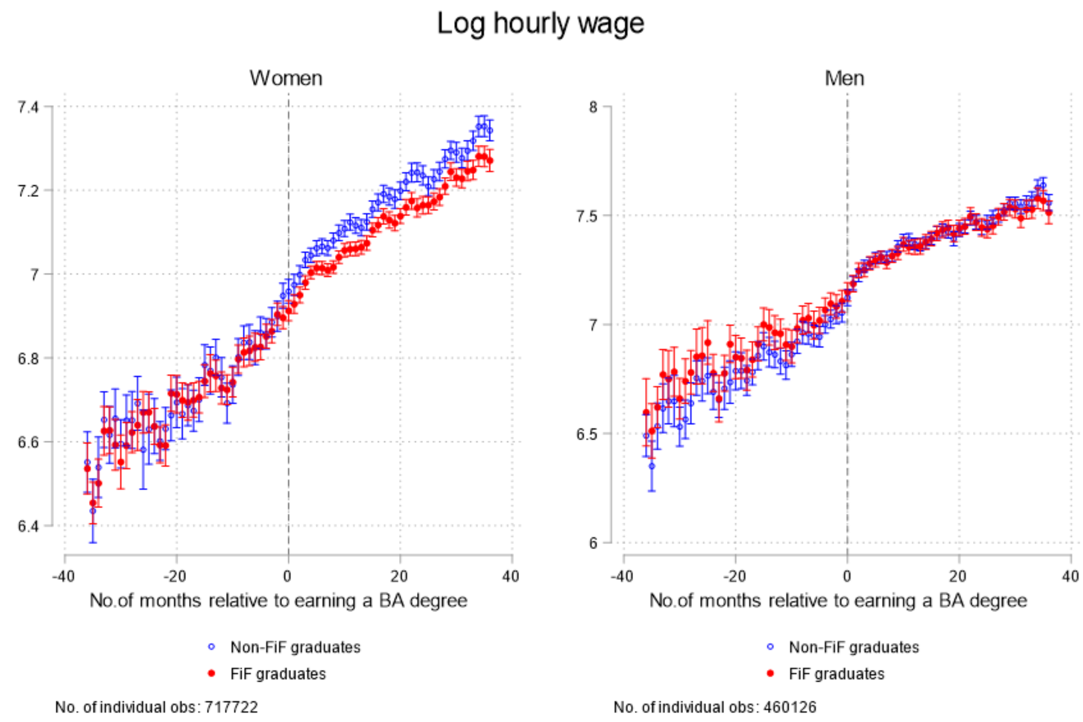
Source: Admin3.

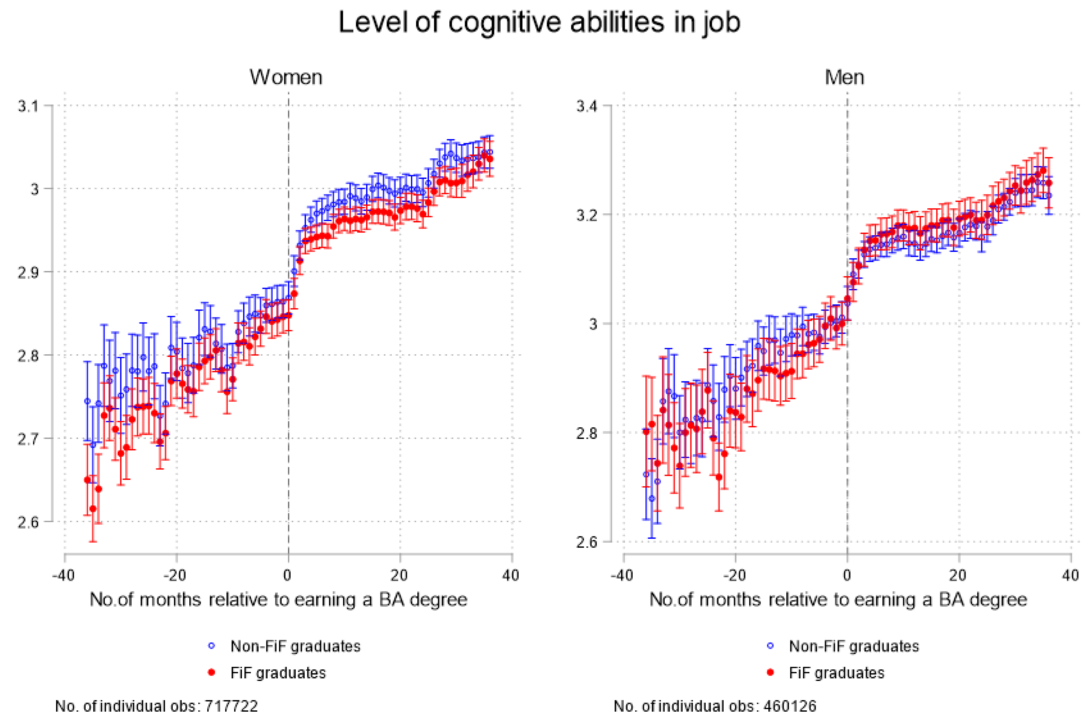
Labor market outcomes

Figure 1 plots the raw labor market outcomes among FiF and non-FiF graduates. Both male and female FiF (BA/BSc) graduates are less likely to work right after graduation compared to non-FiF graduates because FIF graduates are less likely to continue their studies on masters courses. Among female graduates, FiF graduates earn less, and the pay gap emerges right after graduation. Among graduate men, we do not see a raw wage difference by FiF status. Similarly, female FiF graduates tend to work at occupations that require lower levels of cognitive skills, while men do not. Interestingly, when we look at the probability of working in a graduate job (conditional on being employed), there is no clear difference between FiF and non-FiF graduates.

Figure 1: Employment outcomes of graduates by gender and FiF status







Source: Admin3, O*NET.

4. Empirical methods

4.1 Intergenerational educational mobility and gender

First, we investigate intergenerational educational mobility to university by gender. We are interested in whether women or men are more mobile in this respect. We estimate linear probability models on cross-sectional data as

$$y_i = \propto + \beta_1 * female_i + \beta_2 * potential\ FiF_i + \beta_3 * female_i * potential\ FiF_i + \delta * X_i + u_i \quad (1)$$

where y_i stands for one graduation, HE participation or aspirations for university; $female_i$ is a female dummy, $potential\ FiF_i$ is a binary variable indicating that neither parent have a university degree, X_i is a matrix of individual characteristics (low-stake exam test scores in math and reading at grade 10, cohort FE), and u_i is a usual error term. We also estimate the same model separately for men and women.

4.2 FiF wage gap among graduate men and women

We estimate standard Mincer-type wage equations on the subsample of university graduates, using monthly panel data in ages 24-26, pooled and separately by gender as follows:

$$y_{i,t} = \propto + \beta_1 female_i + \beta_2 FiF_i + \beta_3 female_i * FiF_i + \delta_1 X_i + \delta_2 Z_{i,t} + u_i \qquad (2)$$

where $y_{i,t}$ stands for individual labor market outcome variables in month t (log hourly wage, employment, hours worked per week); $female_i$ is a female dummy, FiF_i is a binary variable indicating FiF university graduates, X_i is a matrix of time-invariant individual characteristics (low-stake exam test scores in math and reading at grade 10, cohort FE, subregion FE), $Z_{i,t}$ stands for a matrix of time-variant individual and employment characteristics (age, occupation, industry, firm size) and u_i is a usual error term (robust). Note that as we control for cohort and age, we implicitly control for the time of observation. We also estimate the same model separately for men and women, as before.

4.3 Potential channels

We investigate the role of selection to firms in two steps. First, we re-estimate equation (2) as detailed above, having firm-level characteristics as a dependent variable. In the second step, we extend equation (2) by adding the investigated firm-level characteristic as an additional control variable to the model and look at how the main coefficients of interest (β_2 and β_3) change.

We also investigate the role of firms by extending equation (2) with firm fixed effects, and firm*occupation fixed effects, as well as their interactions with gender. In these later models, we cannot identify the coefficient on female (β_1), only the coefficient of FiF among men (β_2) and among women (β_3).

5. Results

5.1 Intergenerational educational mobility and gender

Table 3 investigates the probability of graduation among men and women who are either potential FiF (i.e. none of their parents have a degree), or at least one of their parents have a degree (this is the baseline group). Women are more likely to graduate then men (Model 1 and 2), while potential FiF young people are less likely to graduate than children of graduate parents. The interaction term of female and potential FiF is negative, meaning that potential FiF women are relatively less likely to graduate than potential FiF men, i.e. women are less mobile in Hungary than men. This is also true on the subsample of those who completed s maturity exam, i.e. fulfilled the requirement of university application (Model 4).

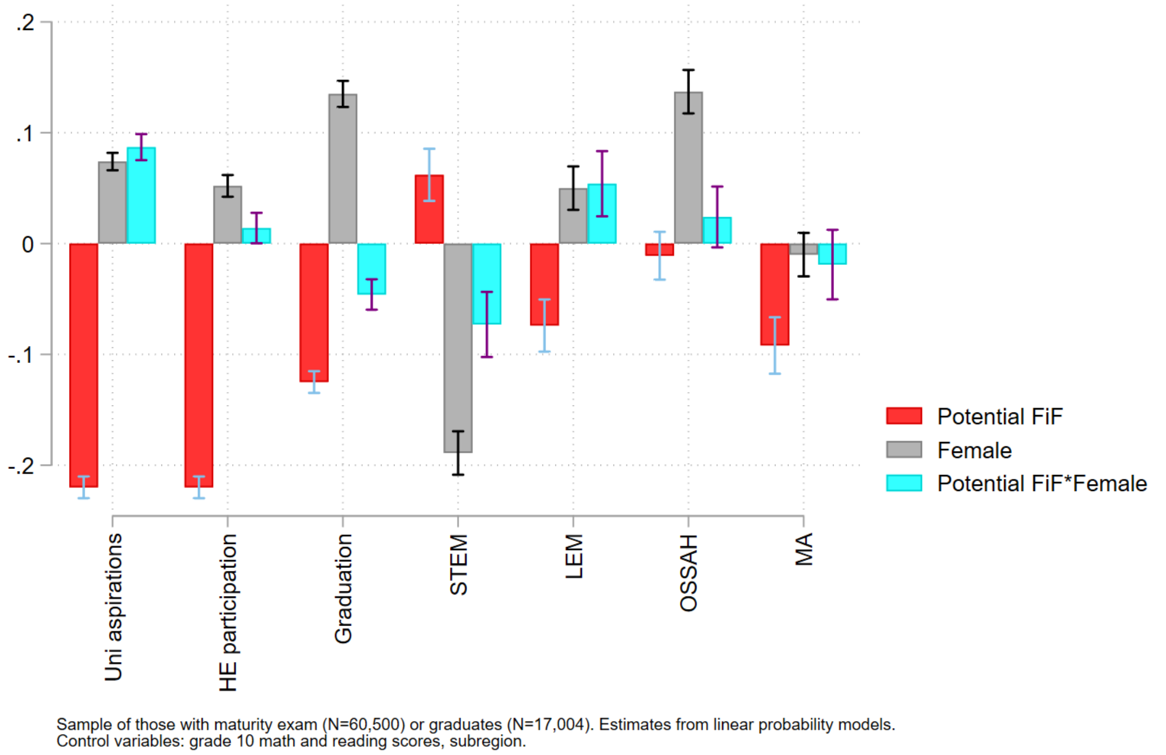
Table 3: The FiF and gender gap in the probability of graduation

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Potential FiF		-0.168***	-0.129***	-0.141***
		(0.001)	(0.002)	(0.002)
Female	0.086***	0.094***	0.149***	0.142***
	(0.001)	(0.001)	(0.002)	(0.002)
Female*potential FiF			-0.079***	-0.049***
			(0.003)	(0.003)
Cohort FE	-0.015***	-0.019***	-0.019***	-0.023***
	(0.001)	(0.001)	(0.001)	(0.002)
Math test score level, grade 10	0.071***	0.063***	0.063***	0.073***
	(0.000)	(0.000)	(0.000)	(0.001)
Reading test score level, grade 10	0.068***	0.059***	0.059***	0.076***
	(0.000)	(0.000)	(0.000)	(0.001)
Constant	-0.405***	-0.219***	-0.247***	-0.369***
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	619,828	503,517	503,517	395,606
R-squared	0.222	0.261	0.263	0.234
Controls	Yes	Yes	Yes	Yes
Sample	All	All	All	Maturity exam

Source: Admin 3. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Further control variables: subregion fixed effects.

Repeating the same exercise for the probability of having university aspirations in academic high school in grade 10, and HE participation reveals that the relative mobility of women to men decreases over these steps (Figure 2). While in high school, potential FiF women lag behind children of graduate parents less than potential FiF men (who have the lowest aspirations for university). In terms of HE participation, potential FiF women still have some advantage over potential FiF men. However, looking at graduation, potential FiF women lag behind children of graduate parents more than potential FiF men. Thus, while in terms of university aspirations and going to university potential FiF women are more mobile than men, once at university, they seem to be less likely to graduate than potential FiF men.

Figure 2: The marginal relationship between HE outcomes and potential FiF*female



Source: Admin3

Looking at university courses among those who graduated, potential FiF women are the least likely to study STEM, but somewhat more likely to study LEM than potential FiF men. Potential FiF status is not related to the probability of studying OSSAH, but negatively associated with the probability of earning an MA.

5.2 The FiF wage gap among graduate men and women

As we have seen in Figure 1, both FiF men and women are more likely to work than non-FiF graduates (Table A1 in Appendix A). The average FiF gap in employment at ages 24-26 is 3.6-3.8 percentage points. Table 4 shows the FiF wage gap among graduates. Model 1 shows that the FiF wage gap on average is 2.6 log points. Looking at the heterogeneity of this gap across genders reveals that conditional on pre-university educational attainment, course, degree type, industry, occupation, firm size, and local labor market fixed effects, the FiF wage gap among graduate women is 4.3 log points larger than among graduate men (Model 4). Looking at the FiF wage gap separately by gender, female FiF graduates earn on average 4.2 log points (4.1%) less than non-FiF graduates, while among men, this difference is 1.1 log points (1.1%).

Table 4: The FiF gap in log hourly wages among graduates at age 24-26 (cohorts born between June 1991 and May 1993)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 4 Women	(6) Model 4 Men
FiF	-0.026***	0.005	0.007*	-0.003	-0.042***	-0.011***
	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
Female	-0.177***	-0.157***	-0.133***	-0.055***		
	(0.002)	(0.003)	(0.003)	(0.004)		
FiF*Female		-0.048***	-0.042***	-0.043***		
		(0.005)	(0.004)	(0.005)		
Constant	7.000***	6.984***	7.049***	6.124***	6.200***	6.181***
	(0.173)	(0.173)	(0.173)	(0.148)	(0.099)	(0.199)
Observations	174,934	174,934	174,934	105,327	60,693	44,634
R-squared	0.140	0.141	0.174	0.290	0.263	0.302
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Degree type			Yes	Yes	Yes	Yes
Industry				Yes	Yes	Yes
Occupation				Yes	Yes	Yes
Firm size				Yes	Yes	Yes
Sample	Graduates	Graduates	Graduates	Graduates working at double-accounting firms	Female graduates working at double-accounting firms	Male graduates working at double-accounting firms

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} – 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

6. Potential channels

6.1 Selection to firms

We investigate selection to firms from three angles. Table 6 shows that average wages are lower at the firms where FiF graduates work, and this difference is about two times as large among women (Column 2) as among men (Column 3). Controlling for firm-level mean wages explains 55% of the FiF wage gap among graduate women, and 73% of the FiF wage gap among graduate men.

Table 6: The gender and FiF gap in the average wage at the firm

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: mean hourly firm wage			Outcome: own log hourly wage, controlling for mean hourly firm wage		
	Total	Women	Men	Total	Women	Men
FiF	-48.853***	-100.962***	-53.619***	0.005	-0.019***	-0.003
	(9.049)	(7.830)	(9.250)	(0.004)	(0.003)	(0.004)
Female	-41.768***			-0.047***		
	(8.625)			(0.003)		
FiF*Female	-53.382***			-0.029***		
	(11.634)			(0.004)		
Constant	-1,524.843***	-519.629***	-1,199.768***	6.383***	6.335***	6.459***
	(122.417)	(143.478)	(109.245)	(0.145)	(0.091)	(0.197)

Observations	105,226	60,628	44,598	104,232	60,058	44,174
R-squared	0.260	0.265	0.305	0.478	0.470	0.472
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} – 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, industry, firm size.

Repeating the same exercise with other firm-level measures as log value added (Table A2 in Appendix A) or log sales per employee (Table A3 in Appendix A) shows a similar picture. Firms where FiF graduates work produce on average lower value added and lower sales revenue per employee than firms where non-FiF graduates work, and these differences explain a fair share of the FiF wage gap. Controlling for log per capita value added explains 88% of the FiF wage penalty on women.

Second, we further investigate the role of firms in the FiF wage gap by extending our baseline model (equation 2) with firm fixed effects in Table A4 and firm*occupation fixed effects in Table A5 in Appendix A. In models with firm fixed effects, our estimated coefficients are very similar to the models that control for firm-level average wages. Among women, the within-firm FiF wage gap is 2.8 log points (Model 2), among men, it is 1.0 log points (Model 3). When we pool the data of men and women together and apply female*firm fixed effects (i.e., assume that firm FE's differ by gender within firms), the FiF gap among men reduces to close to zero (insignificant 0.6 log points), while among women, it is similar to the earlier estimates (2.3 log points). Introducing firm*occupation fixed effects allows us to estimate gender*FiF gaps within firms-and-occupations. Among those working in the same occupation at the same firm, the estimated FiF gaps are very similar for men and women at 1.8 and 1.9 log points (Model 2 and 3).

Table 7: The gender and FiF gap in wages: the role firm-level wage premia

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: firm-level wage premia (normalized)			Outcome: own log hourly wage, controlling for firm wage premia		
	Total	Women	Men	Total	Women	Men
FiF	0.001	-0.014***	-0.002	0.010**	-0.007*	0.003
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
Female	0.011***			-0.067***		
	(0.004)			(0.004)		
FiF*Female	-0.021***			-0.020***		
	(0.005)			(0.006)		
Constant	-0.147***	-0.169***	-0.122***	6.922***	6.963***	7.003***
	(0.056)	(0.058)	(0.040)	(0.101)	(0.173)	(0.099)
Observations	49,884	25,551	24,333	49,688	25,446	24,242
R-squared	0.232	0.306	0.276	0.460	0.466	0.464
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, firm size.

Third, instead of looking at firm-level average wages, we estimate firm-level wage premia (firm FE's) exploiting the total sample of firms and individuals that the database covers. Table 7 shows that firm-level wage premia are lower at firms where FiF graduate women work (Column 2), while for men, there is no such difference (Column 1). When we control for the firm-specific wage premia, it explains 83% of the FiF wage gap among women and 73% of the FiF wage gap among men.

Table 8: The gender and FiF gap in wages: the role of the gender gap in firm-level wage premia

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: gender difference in firm-level wage premia (men-women)			Outcome: own log hourly wage, controlling for the gender gap in firm wage premia		
	Total	Women	Men	Total	Women	Men
FiF	0.013***	-0.001	0.008***	0.005	-0.019***	-0.003
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Female	0.024***			-0.070***		
	(0.002)			(0.005)		
FiF*Female	-0.018***			-0.027***		
	(0.003)			(0.007)		
Constant	0.141***	0.245***	0.017	6.714***	6.690***	6.894***
	(0.029)	(0.040)	(0.022)	(0.115)	(0.196)	(0.105)
Observations	48,535	24,977	23,558	48,351	24,872	23,479
R-squared	0.125	0.204	0.189	0.323	0.325	0.329
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, firm size.

Table 8 investigates the role of gender differences in the estimated firm-level wage premia (men-women). In the pooled model (Column 1), both FiF men and women are somewhat more likely to work at firms with higher gender imbalance in wage premia. Once investigated separately by gender in Column 2 and 3, these associations decrease towards zero. Controlling for gender gap in wage premia explains 53% if the female and 73% of the male FiF gap.

6.2 Selection to occupations

Table 9 shows that FiF women tend to work in occupations that require lower cognitive abilities on average that occupations taken by non-FiF graduates (Model 2). Controlling for the occupation-level importance of cognitive abilities explains 21% of the female FiF wage penalty. As we have seen before, the importance of cognitive abilities does not differ between the occupations of FiF and non-FiF male graduates.

Table 9: The importance of cognitive abilities in jobs

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: the importance of cognitive abilities for the job			Outcome: own log hourly wage, controlling for the importance of cognitive abilities for the job		
	Total	Women	Men	Total	Women	Men
FiF	-0.013	-0.030***	-0.018	-0.012	-0.032***	-0.010
	(0.015)	(0.011)	(0.014)	(0.011)	(0.010)	(0.012)
Female	-0.068***			-0.031***		
	(0.014)			(0.010)		
FiF*Female	-0.020			-0.016		
	(0.018)			(0.014)		
Constant	2.969***	2.995***	3.076***	6.426***	6.356***	6.107***
	(0.081)	(0.190)	(0.108)	(0.079)	(0.208)	(0.132)
Observations	94,691	57,965	36,664	93,004	56,919	36,025
R-squared	0.798	0.814	0.853	0.731	0.742	0.749
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors clustered by subregion in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, firm FE.

6.3 The role of fertility and female-friendly firms

As we have shown, FiF graduates (and especially women), tend to work at “worse” firms than non-FiF graduates. This pattern could occur because of (at least) three reasons. First, even though they apply at the same rate, there may be a “class ceiling” (Friedman and Laurison 2019) where “good” firms might be less likely to accept FiF candidates than “worse” firms. Second, FiF graduates have lower productivity, hence they are selected to worse firms. These first two potential channels we cannot test (for now). Third, FiF graduates may value non-pecuniary elements of jobs more than non-FiF graduates. We measure how “good’ firms are via financial measures (wage, value added, wage premia), but it might be that FiF graduates value something else more. For example, FiF women may care more about how “female friendly” a firm is and might accept a lower wage in exchange for more flexibility. In this subsection, we test this latter hypothesis.

At age 24-26, FiF graduate women are 0.3 log points less likely to have children, than non-FiF female graduates (Table 10, Column 1). Controlling for having children does not change the FiF gap among women, although as expected, women who have children earn substantially less than women who have not (Column 2). The child penalty does not differ between FiF and non-FiF graduates (Column 3).

Graduate women who work at firms where the share of women is larger, earn less (Column 4). However, if they are FiF, this negative association is somewhat smaller. Interpreting this result together with the somewhat larger FiF wage gap in Model 4 suggests that female FiF wage penalty is smaller at firms where the share of women is larger. Lastly, we look at the share of mothers among women working at the same firms in Column 5. Women on average earn more at firms where the share of mothers is larger, but this return is smaller for FiF women and for women who have children.

Table 10: The FiF gap among women: the role children

	(1)	(2)	(3)	(4)	(5)
	Outcome: has child	Outcome: log hourly wages			
FiF graduate	-0.003***	-0.042***	-0.042***	-0.067***	-0.013**
	(0.001)	(0.003)	(0.003)	(0.008)	(0.007)
Has child		-0.292***	-0.267***	-0.365***	0.183**
		(0.033)	(0.039)	(0.087)	(0.074)
FiF graduate*has child			-0.074	0.093	-0.168
			(0.070)	(0.287)	(0.184)
Share of women at the firm				-0.222***	
				(0.011)	
FiF graduate*share of women at the firm				0.044***	
				(0.014)	
Has child*share of women at the firm				0.193	
				(0.151)	
FiF graduate*has child*share of women at the firm				-0.269	
				(0.450)	
Share of women with children					0.458***
					(0.016)
FiF graduate*share of women with children					-0.124***
					(0.022)
Has child*share of women with children					-1.373***
					(0.143)
FiF graduate*has child*share of women with children					0.399
					(0.455)
Constant	0.018***	6.203***	6.204***	6.330***	6.045***
	(0.006)	(0.099)	(0.099)	(0.100)	(0.100)
Observations	61,517	60,693	60,693	60,693	60,693
R-squared	0.054	0.264	0.264	0.272	0.282

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} - 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, firm size. Sample of graduate women.

Model 1 and 2 in Table A11 in Appendix A suggest that the female FiF penalty is larger at firms where the probability that a female employee will come back to the firm to the same occupation after maternity leave is higher. Women earn less on average at firms where the wage drop after coming back from maternity leave is lower (or the wage increase is higher), but this is not related to being FiF (Model 3). Lastly, women earn less at firms where women on average go on a longer maternity leave, and the FiF wage penalty is also higher at these firm for FiF mothers.

FiF women tend to work at firms where the share of women is lower (Table A6), the share of mothers among women is a somewhat higher (Table A7), the probability that women will come back to the same firm and occupation is lower (Table A8), women's relative wage after maternity leave compared to before maternity leave is higher (Table A9), and maternity leave is longer (Table A10) than at firm where non-FiF graduates work. However, these measures only explain a small share of the FiF wage penalty among women.

7. Discussion

This paper investigated the first in family wage gap among graduate men and women. We showed that in a usual Mincer-type framework, FiF graduate women earn on average 4.1% less than graduate women whose parents are graduates, while among men, this difference is 1.1%.

Exploiting linked employer-employee data as well as the cognitive ability content of jobs, we investigated the role of firms and occupations behind the FiF wage penalty. We find that FiF graduates, especially women, tend to work at somewhat "worse" firms than graduate children of graduate parents. Firms where FiF graduates work pay on average lower wages, and produce less value added and sales revenues per employee

Interestingly, when we look at the FiF wage gap among those who work at the same firms in the same occupations (i.e., we apply firm*occupation FE's), the FiF gap is of the same magnitude for men and women, at 1.8-1.9 log points (1.8-1.9%). This result explains an earlier finding of the literature. Decomposing the FiF wage gap to endowment differences and differential returns to these endowments using an Oaxaca-decomposition, Adamecz-Völgyi, Henderson, and Shure (2022) find that FiF men and women have similar characteristics that contribute to lower wages. For example, they both "undermatch" on the labor market by working in jobs that would not require their highest degree. However, FiF men somehow compensate this phenomenon and do not suffer a wage penalty (while FiF women do). Our current results show that within the same occupation-and-firm, the FiF wage gap is basically the same for both men and women. Thus, men "compensate" for their FiF-related disadvantage by sorting to better paying jobs (occupation-and-firm pairs), even though they might still undermatch. Looking at sorting to firms or occupations separately would not explain this puzzle.

Looking at the occupation-level importance of cognitive abilities, FiF graduate women tend to work in occupation that require lower levels of abilities than non-FiF graduate women. Among men, we do not see this difference.

Our analysis is not without caveats. First, due to the limitations of the data, we only observe graduates up until age 25 on average. This is a very early age to look at labor market outcomes, although the literature suggests that labor market entry defines one's long-term labor market outcomes. Second, we cannot identify the causal effects of being a FiF graduate; parents and children are not allocated randomly to each other, so all we can investigate are (conditional) correlations. Still, as we observe math and reading test scores from high school, the details of young peoples' HE degrees, as well as firm characteristics and firm/occupation, we are able to control for a rich range of characteristics. We thus believe that the statistical associations that we find are meaningful.

From a policy point of view, our results confirm that HE degrees are not automatic equalizers. Our results suggest that FiF graduates might need more career advice at universities to make better informed decisions about their labor market entry. This is especially true for FiF women, who may be less likely seek out networks or self-promote themselves (Exley and Kessler 2022).

References

Adamecz-Völgyi, Anna, Morag Henderson, and Nikki Shure. 2022. "The Labor Market Returns to 'First-in-Family' University Graduates." *Journal of Population Economics*, May. https://doi.org/10.1007/s00148-022-00908-y.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *The Quarterly Journal of Economics* 131 (2): 633–86. https://doi.org/10.1093/qje/qjv038.

European Commission. 2017. "A Renewed EU Agenda for Higher Education." Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions {COM(2017) 247 final}. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017SC0164&from=EN.

Exley, Christine L, and Judd B Kessler. 2022. "The Gender Gap in Self-Promotion." *The Quarterly Journal of Economics*, January, qjac003. https://doi.org/10.1093/qje/qjac003.

Friedman, Sam, and Daniel Laurison. 2019. *The Class Ceiling: Why It Pays to Be Privileged*. 1st ed. Policy Press. https://doi.org/10.2307/j.ctv5zftbj.

Henderson, Morag, Nikki Shure, and Anna Adamecz-Völgyi. 2020. "Moving on up: 'First in Family' University Graduates in England." *Oxford Review of Education*. https://doi.org/10.1080/03054985.2020.1784714.

Manzoni, Anna, and Jessi Streib. 2019. "The Equalizing Power of a College Degree for First-Generation College Students: Disparities Across Institutions, Majors, and Achievement Levels." *Research in Higher Education*, no. 60: 577–605.

Nunez, Anne-Marie, and Stephanie Cuccaro-Alamin. 1998. "First-Generation Students: Undergraduates Whose Parents Never Enrolled in Postsecondary Education." Statistical Analysis Report. National Center for Education Statistics. https://www.researchgate.net/publication/242182970_Students_Whose_Parents_Did_Not_Go_to_College_Postsecondary_Access_Persistence_and_Attainment.

Oreopoulos, Philip, and Uros Petronijevic. 2013. "Making College Worth It: A Review of Research on the Returns to Higher Education." Working Paper 19053. National Bureau of Economic Research. https://doi.org/10.3386/w19053.

Sebők, Anna. 2019. "The Panel of Linked Administrative Data of CERS Databank: A KRTK Adatbank Kapcsolt Allamigazgatási Paneladatbázisa." 2. Budapest Working Papers on the Labour Market. Institute of Economics, Centre for Economic and Regional Studies.

Thomas, Scott L., and Liang Zhang. 2005. "Post-Baccalaureate Wage Growth within Four Years of Graduation: The Effects of College Quality and College Major." *Research in Higher Education* 46 (4): 437–59. https://doi.org/10.1007/s11162-005-2969-y.

Appendix A

Table A1: The gender and FiF gap in the probability of employment

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 3 Women	(5) Model 3 Men
FiF	0.049***	0.040***	0.034***	0.038***	0.036***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)
Female	0.038***	0.033***	0.029***		
	(0.002)	(0.002)	(0.002)		
FiF*Female		0.013***	0.005		
		(0.004)	(0.004)		
Constant	-0.073	-0.068	0.113*	0.028	0.223***
	(0.055)	(0.055)	(0.059)	(0.128)	(0.057)
Observations	317,055	317,055	317,055	198,029	119,026
R-squared	0.048	0.048	0.082	0.085	0.096
Controls	Yes	Yes	Yes	Yes	Yes
Degree type			Yes	Yes	Yes
Sample	Graduates	Graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables: age, cohort, subregion, grade 10 math and reading scores, subregion FE.

Table A2: The gender and FiF gap in log hourly wages: the role of per capita log value added

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: log value added per employee			Outcome: own log hourly wage		
	Total	Women	Men	Total	Women	Men
FiF	-0.076***	-0.110***	-0.082***	0.019***	-0.005	0.012**
	(0.014)	(0.015)	(0.014)	(0.005)	(0.005)	(0.005)
Female	-0.030**			-0.055***		
	(0.014)			(0.005)		
FiF*Female	-0.058***			-0.027***		
	(0.019)			(0.006)		
Constant	8.358***	9.550***	6.924***	5.902***	5.874***	6.061***
	(0.266)	(0.335)	(0.206)	(0.102)	(0.176)	(0.104)
Observations	49,909	25,562	24,347	49,713	25,457	24,256
R-squared	0.326	0.369	0.392	0.377	0.373	0.391
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Degree type	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Firm size	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} – 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A3: The gender and FiF gap in log hourly wages: the role of per capita log sales revenues

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: log sales revenues per employee			Outcome: own log hourly wage		
	Total	Women	Men	Total	Women	Men
	-0.115***	-0.128***	-0.134***	0.017***	-0.031***	0.006
FiF	(0.012)	(0.011)	(0.012)	(0.004)	(0.004)	(0.004)
	-0.034***			-0.041***		
Female	(0.012)			(0.004)		
	-0.014			-0.053***		
FiF*Female	(0.016)			(0.006)		
	10.161***	10.411***	9.797***	6.198***	6.234***	6.152***
Constant	(0.160)	(0.213)	(0.138)	(0.079)	(0.115)	(0.100)
	76,814	41,767	35,047	76,311	41,527	34,784
Observations	0.286	0.351	0.301	0.305	0.296	0.319
R-squared	-0.115***	-0.128***	-0.134***	0.017***	-0.031***	0.006
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Degree type	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Firm size	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A4: The FiF gap in log hourly wages – models with firm fixed effects

	(1) Model 1 Firm FE	(2) Model 1 Firm FE Women	(3) Model 1 Firm FE Men	(4) Model 1 Firm*female FE
FiF	-0.005	-0.028***	-0.010**	-0.006
	(0.004)	(0.004)	(0.005)	(0.005)
Female	-0.032***			
	(0.004)			
FiF*Female	-0.022***			-0.023***
	(0.005)			(0.006)
Constant	7.150***	7.060***	7.293***	7.173***
	(0.031)	(0.103)	(0.043)	(0.034)
Observations	104,939	60,409	44,468	104,877
R-squared	0.742	0.749	0.755	0.755
Controls	Yes	Yes	Yes	Yes
Degree type	Yes	Yes	Yes	Yes
Industry	No	No	No	No
Occupation	Yes	Yes	Yes	Yes
Firm size (yearly)				
Firm FE	Yes	Yes	Yes	No
Firm*Female FE	No	No	No	Yes
Sample	Graduates	Female graduates	Male graduates	Graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A5: The FiF gap in log hourly wages – models with firm*occupation fixed effects

	(1)	(2)	(3)	(4)
	Model 1 Firm*Occ FE	Model 1 Firm*Occ FE Women	Model 1 Firm*Occ FE Men	Model 1 Firm*Occ*Female FE
FiF	-0.006	-0.018***	-0.019***	-0.012**
	(0.004)	(0.004)	(0.005)	(0.005)
Female	-0.036***			
	(0.004)			
FiF*Female	-0.024***			-0.013**
	(0.006)			(0.006)
Constant	7.274***	7.340***	7.454***	7.266***
	(0.035)	(0.118)	(0.047)	(0.038)
Observations	104,805	60,318	44,418	104,736
R-squared	0.786	0.794	0.791	0.798
Controls	Yes	Yes	Yes	Yes
Degree type	Yes	Yes	Yes	Yes
Industry	No	No	No	No
Occupation	Yes	Yes	Yes	Yes
Firm*Occupation FE	Yes	Yes	Yes	No
Firm*Occupation*Female FE	No	No	No	Yes
Sample	Graduates	Female graduates	Male graduates	Graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A6: The FiF gap in log hourly wages – the role of the share of women at the firm

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: share of women at the firm			Outcome: own log hourly wage		
		Women	Men		Women	Men
FiF	0.008***	-0.003*	0.006***	-0.003	-0.042***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)
Female	0.103***			-0.068***		
	(0.002)			(0.004)		
FiF*Female	-0.018***			-0.041***		
	(0.002)			(0.005)		
Constant	0.507***	0.611***	0.359***	6.192***	6.317***	6.225***
	(0.033)	(0.039)	(0.025)	(0.149)	(0.100)	(0.199)
Observations	103,573	61,517	42,056	102,402	60,693	41,709
R-squared	0.377	0.317	0.352	0.304	0.271	0.311
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A7: The FiF gap in log hourly wages – the role of the share of mothers among women at the firm

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: share of mothers among women at the firm			Outcome: own log hourly wage		
		Women	Men		Women	Men
	0.010***	0.004***	0.008***	-0.005	-0.043***	-0.011**
FiF	(0.002)	(0.001)	(0.002)	(0.004)	(0.003)	(0.004)
	-0.036***			-0.074***		
Female	(0.002)			(0.004)		
	-0.006**			-0.038***		
FiF*Female	(0.002)			(0.005)		
	0.295***	0.625***	-0.129***	6.139***	6.078***	6.241***
Constant	(0.099)	(0.025)	(0.024)	(0.148)	(0.100)	(0.199)
	103,573	61,517	42,056	102,402	60,693	41,709
Observations	0.145	0.188	0.136	0.305	0.278	0.311
R-squared	0.010***	0.004***	0.008***	-0.005	-0.043***	-0.011**
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} – 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A8: The FiF gap in log hourly wages – the role of the probability that mothers will come back to the same firm and occupation after maternity leave

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: the probability that mothers will come back to the same firm and occupation after maternity leave at the firm			Outcome: own log hourly wage		
		Women	Men		Women	Men
FiF	0.004**	-0.006***	0.002	-0.005	-0.038***	-0.011***
	(0.002)	(0.001)	(0.002)	(0.004)	(0.003)	(0.004)
Female	-0.021***			-0.073***		
	(0.001)			(0.004)		
FiF*Female	-0.011***			-0.035***		
	(0.002)			(0.005)		
Constant	0.223*	0.701***	-0.267***	6.177***	6.000***	6.277***
	(0.127)	(0.022)	(0.022)	(0.148)	(0.098)	(0.199)
Observations	103,573	61,517	42,056	102,402	60,693	41,709
R-squared	0.094	0.128	0.098	0.313	0.288	0.313
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: $100*(e^{beta} – 1)$, where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A9: The FiF gap in log hourly wages – the role of relative wages after and before maternity leave (after/before)

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: relative wages after and before maternity leave (after/before)			Outcome: own log hourly wage		
		Women	Men		Women	Men
FiF	-0.001	0.013***	-0.003	-0.011**	-0.047***	-0.018***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Female	-0.021***			-0.065***		
	(0.002)			(0.004)		
FiF*Female	0.015***			-0.039***		
	(0.002)			(0.005)		
Constant	1.536***	1.334***	0.954***	6.506***	6.510***	6.802***
	(0.023)	(0.029)	(0.029)	(0.070)	(0.104)	(0.098)
Observations	83,988	48,771	35,217	83,556	48,486	35,070
R-squared	0.085	0.119	0.107	0.308	0.279	0.319
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A10: The FiF gap in log hourly wages – the role of length of maternity leave

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: length of maternity leave at the firm			Outcome: own log hourly wage		
		Women	Men		Women	Men
FiF	-0.140	0.732***	-0.281**	-0.011***	-0.044***	-0.019***
	(0.126)	(0.117)	(0.129)	(0.004)	(0.004)	(0.004)
Female	-0.382***			-0.065***		
	(0.117)			(0.004)		
FiF*Female	0.929***			-0.036***		
	(0.165)			(0.005)		
Constant	87.538***	77.794***	86.785***	6.845***	6.827***	7.152***
	(1.518)	(2.007)	(1.697)	(0.070)	(0.103)	(0.100)
Observations	83,988	48,771	35,217	83,556	48,486	35,070
R-squared	0.090	0.119	0.156	0.323	0.295	0.334
Sample	Graduates	Female graduates	Male graduates	Graduates	Female graduates	Male graduates

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores.

Table A11: The heterogeneity of the FiF wage gap among women: the role children

	(1)	(2)	(3)	(4)
	Outcome variable: log hourly wages			
FiF graduate	-0.028***	-0.028***	-0.071***	-0.053***
	(0.006)	(0.006)	(0.025)	(0.011)
Has child	0.078	0.078	-0.545	-0.463***
	(0.061)	(0.061)	(0.407)	(0.097)
FiF graduate*has child	-0.116	-0.116	-0.106	0.940***
	(0.138)	(0.138)	(0.722)	(0.171)
Average probability of coming back to the same occupation at the same firm after maternity leave	0.587***			
	(0.020)			
FiF graduate*average probability of coming back to the same occupation at the same firm after maternity leave	-0.054**			
	(0.027)			
Has child*average probability of coming back to the same occupation at the same firm after maternity leave	-1.319***			
	(0.142)			
FiF graduate*has child*average probability of coming back to the same occupation at the same firm after maternity leave	0.100			
	(0.510)			
Average probability of coming back to the same firm after maternity leave		0.587***		
		(0.020)		
FiF graduate*average probability of coming back tot he same firm after maternity leave		-0.054**		
		(0.027)		
Has child*average probability of coming back tot he same firm after maternity leave		-1.319***		
		(0.142)		
FiF graduate*has child*average probability of coming back to the same firm after maternity leave		0.100		
		(0.510)		
Wage after leave over before			-0.073***	
			(0.014)	
FiF graduate*wage after leave over before			0.021	
			(0.021)	
Has child* wage after leave over before			0.202	
			(0.325)	
FiF graduate*has child* wage after leave over before			0.122	
			(0.593)	
Length of maternity leave				-0.005***
				(0.000)
FiF*length of maternity leave				0.000
				(0.449)
Has child*length of maternity leave				0.005**
				(0.002)
FiF graduate*has child*length of maternity leave				-0.027***
				(0.005)
Constant	5.988***	5.988***	6.524***	6.837***
	(0.098)	(0.098)	(0.105)	(0.103)
Observations	60,693	60,693	48,486	48,486
R-squared	0.292	0.292	0.281	0.298

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For log hourly wage, coefficients are in log points and may be transformed to percentages through the following transformation: 100*(e^{beta} – 1), where beta is the estimated coefficient. Control variables: age, cohort, subregion, grade 10 math and reading scores, degree type, firm size. Sample of graduate women.

Appendix B

Table B1: Abilities in O*NET 20.1 used in the paper

Element ID	Element Name	Description
1	Worker Characteristics	Worker Characteristics
1.A	Abilities	Enduring attributes of the individual that influence performance
1.A.1	Cognitive Abilities	Abilities that influence the acquisition and application of knowledge in problem solving
1.A.1.a	Verbal Abilities	Abilities that influence the acquisition and application of verbal information in problem solving
1.A.1.a.1	Oral Comprehension	The ability to listen to and understand information and ideas presented through spoken words and sentences.
1.A.1.a.2	Written Comprehension	The ability to read and understand information and ideas presented in writing.
1.A.1.a.3	Oral Expression	The ability to communicate information and ideas in speaking so others will understand.
1.A.1.a.4	Written Expression	The ability to communicate information and ideas in writing so others will understand.
1.A.1.b	Idea Generation and Reasoning Abilities	Abilities that influence the application and manipulation of information in problem solving
1.A.1.b.1	Fluency of Ideas	The ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).
1.A.1.b.2	Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
1.A.1.b.3	Problem Sensitivity	The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.
1.A.1.b.4	Deductive Reasoning	The ability to apply general rules to specific problems to produce answers that make sense.
1.A.1.b.5	Inductive Reasoning	The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).
1.A.1.b.6	Information Ordering	The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).
1.A.1.b.7	Category Flexibility	The ability to generate or use different sets of rules for combining or grouping things in different ways.
1.A.1.c	Quantitative Abilities	Abilities that influence the solution of problems involving mathematical relationships
1.A.1.c.1	Mathematical Reasoning	The ability to choose the right mathematical methods or formulas to solve a problem.
1.A.1.c.2	Number Facility	The ability to add, subtract, multiply, or divide quickly and correctly.
1.A.1.d	Memory	Abilities related to the recall of available information

1.A.1.d.1	Memorization	The ability to remember information such as words, numbers, pictures, and procedures.
1.A.1.e	Perceptual Abilities	Abilities related to the acquisition and organization of visual information
1.A.1.e.1	Speed of Closure	The ability to quickly make sense of, combine, and organize information into meaningful patterns.
1.A.1.e.2	Flexibility of Closure	The ability to identify or detect a known pattern (a figure, object, word, or sound) that is hidden in other distracting material.
1.A.1.e.3	Perceptual Speed	The ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns. The things to be compared may be presented at the same time or one after the other. This ability also includes comparing a presented object with a remembered object.
1.A.1.f	Spatial Abilities	Abilities related to the manipulation and organization of spatial information
1.A.1.f.1	Spatial Orientation	The ability to know your location in relation to the environment or to know where other objects are in relation to you.
1.A.1.f.2	Visualization	The ability to imagine how something will look after it is moved around or when its parts are moved or rearranged.
1.A.1.g	Attentiveness	Abilities related to application of attention
1.A.1.g.1	Selective Attention	The ability to concentrate on a task over a period of time without being distracted.
1.A.1.g.2	Time Sharing	The ability to shift back and forth between two or more activities or sources of information (such as speech, sounds, touch, or other sources).

Source: https://www.onetcenter.org/dictionary/20.1/excel/content_model_reference.html