CPB Netherlands Bureau for Economic Policy Analysis

Skill up or get left behind? Digital skills and labor market outcomes in the Netherlands

We use a unique dataset combining the 2012 OECD Survey of Adult Skills with register data on labor market outcomes in 2012-2019. People with higher digital skills have four to six percent higher hourly wages and are ten percent more likely to be employed. The association between digital skills and both wage and employment is stable over time suggesting a long-lasting relationship with labor market outcomes.

We find that individuals with low digital skills are generally older, lower educated and more often female.

CPB Discussion Paper

Mariëlle Non, Milena Dinkova, Benjamin Dahmen

February 2021

Doi: https://doi.org/10.34932/eyzo-4g11

Skill up or get left behind? Digital skills and labor market outcomes in the Netherlands

Mariëlle Non[1]*, Milena Dinkova[1,2] and Benjamin Dahmen[1]
[1]CPB Netherlands Bureau for Economic Policy Analysis, The Netherlands
[2]Utrecht University School of Economics, The Netherlands

February 1, 2021

Abstract

This paper studies the demographic characteristics of people with different digital skill levels and relates digital skills to labor market outcomes in the Netherlands. We use a unique dataset combining the 2012 OECD Survey of Adult Skills with register data on labor market outcomes in 2012-2019. Individuals with low digital skills are generally older, lower educated and more often female. We find that a one standard deviation increase in digital skills is associated with a four to six percent increase in wage. However, there is no significant wage difference between individuals with no skills and those with below basic skills. Persons with at least basic skills are about ten percent more likely to be employed compared to persons with no digital skills. This is mainly due to a higher labor force participation of those with at least basic skills. We find no relation between digital skills and job stability. The association between digital skills and both wage and employment is stable over time, suggesting that skills in 2012 have a long-lasting association with labor market outcomes.

JEL Classification: J24, J31, C21, C83
Keywords: Digital skills, labor market, wage, employment, labor force participation, PIAAC

Acknowledgements: Anne Marieke Braam and Rinske Windig provided much appreciated research assistance. We thank seminar participants at ROA Maastricht (the Netherlands) and the KVS new paper session 2020 for their valuable comments and suggestions. We also would like to thank our CPB-colleagues, in particular Jochem Zweerink and Bas Scheer, for their help. The results in this paper are based on calculations using non-public microdata from Statistics Netherlands.

*Corresponding author: m.c.non@cpb.nl.

1 Introduction

During the last decades, the economy has rapidly become ICT-driven. Not only has the way of working changed, but so has the set of skills required for numerous jobs. These skills concern practical knowledge on how to use digital tools like email and the internet, but also knowledge on how to structure a problem and to evaluate large amounts of digital information (Levy, 2010). For example, car mechanics nowadays regularly use a computer system to diagnose cars. For this task, the mechanic needs basic computer skills to access the system. But (s)he also needs to interpret the information in the system and determine which facts are relevant for the case at hand and which repairs would solve the issue.

Recently, policymakers have expressed concern that certain groups in society are lagging behind when it comes to digital skills and that this affects their opportunities both at the labor market and in society in general. Based on the 2020 Digital Economy and Society Index (DESI), the European Commission concludes that more progress in digital skills is needed, because 42 percent of the EU population still lacks at least basic digital skills (European Commission, 2019c). The 2020 European Skills agenda underlines the importance of skills for jobs in the digital era (European Commission, 2019b). Even in the Netherlands, one of the three most digitalized countries in the European Union (European Commission, 2019a) and the focus of our study, a substantial group of people still lacks basic digital skills. Several recent Dutch policy reports express concerns about this. For instance, a report for the Dutch Rathenau Institute mentions that lower educated people have less digital skills, which is likely to increase inequality (Schel & Vennekens, 2020).[1] The Dutch Ministry of Economic Affairs and Climate also states that life-long learning and assisting people with low digital skills to adapt to the changing labor market is essential to ensure no-one is left behind (Ministry of Economic Affairs and Climate, 2019).

Despite the concerns, there is only very limited empirical evidence on the role of digital skills in the labor market. In this paper, we provide more insight in this topic. We address two questions: 1) How do digital skills vary with demographic characteristics and 2) What is the relation between digital skills and labor market outcomes? We use a unique dataset that combines questionnaire data from the OECD with register data from the Netherlands. The 2012 OECD Survey of Adult Skills (PIAAC) contains an extensive measure of digital skills obtained from assignments in a digital environment. We combine

[1] The Rathenau Institute is a think-tank on the societal aspects of technology, science and innovation.

these data with detailed register data on labor market outcomes in 2012-2019.

To our knowledge, only Falck et al. (2020) explicitly study the returns to digital skills. The authors exploit exogenous variation in internet availability and use this as an instrument for the digital skills as measured in the PIAAC questionnaire. They find that a one standard deviation increase in digital skills increases wages by 24 (internationally) to 31 (for Germany) percent. In our study, we find that a one standard deviation increase in digital skills is associated with a four to six percent increase in hourly wage, dependent on the specification we use. However, a direct comparison with Falck et al. (2020) is difficult as we focus on the Netherlands and cannot claim causality. We cannot use the same identification strategy because internet availability barely varies within the Netherlands. Lane & Conlon (2016) study the impact of digital skills on wage and employment as part of a larger study that also considers the impact of literacy and numeracy skills. Again, this study is based on the PIAAC questionnaire. They find a ten to fifteen percent wage premium for those with basic digital skills, compared to those with below-basic skills. Furthermore, they find a five to eight percent employment effect when comparing those who have no computer experience at all with those with below-basic skills. Again, our findings on wage are more modest in size, but our findings on employment are quite similar.

In this paper, we extend the literature in several directions. First, our rich register data allow us to consider not only wage and employment, but also labor force participation and job stability, topics that are at the core of numerous policy discussions. Our analysis shows that the lower employment rate of those with no or below-basic digital skills is largely due to a lower labor force participation and that there is no correlation between digital skills and unemployment spells or the probability of having a fixed contract. Second, we study labor market outcomes over the period 2012-2019, whereas previous studies focus on a single year. On one hand, one might expect a decreasing association between the skills as measured in 2012 and subsequent labor market outcomes, because over time people may lose or gain skills. On the other hand, digital skills probably have become more important over time. Moreover, a person with above average digital skills in 2012 is likely to have a high-skilled job, which in turn allows a continuous updating of skills. In our research, we do not find any time effects, which means that the relationship between digital skills in 2012 and labor market outcomes in 2012-2019 is persistent. Third, by using register data, we are able to incorporate more accurate information on an individual's wage and employment status than when using self-reported labor market outcomes from the PIAAC questionnaire. As Gauly et al. (2019) show in the case of Germany, using self-reported outcome variables from PIAAC can cause biased outcomes of econometric analyses. Our analysis confirms this finding from Gauly et al. (2019). Fourth, we add to Falck et al. (2020) and follow Lane & Conlon (2016) by explicitly including those with no computer skills in our analysis. We find that the wage of this group does not differ from those with below-basic digital skills, but that their labor force participation is substantially lower. Finally, the PIAAC skill measurement in the form of plausible values requires a specific estimation technique. Although OLS regression, such as used by Falck et al. (2020), gives unbiased estimates, the OLS standard errors are biased downwards. In other words, OLS regression leads to inflated significance. We correctly estimate the standard errors using the Repest package in Stata.

This paper has the following structure. In Section 2 we provide an overview of related literature and in Section 3 we describe our dataset. Section 4 briefly outlines the methodology and Section 5 contains descriptive statistics. We proceed with the results on wages in Section 6 and the results on employment in Section 7. Section 8 gives the results when we use self-reported labor market outcomes from PIAAC. At last, we offer a brief discussion with conclusions in Section 9.

2 Related literature

Several scholars and statisticians have examined the distribution of digital skills over the population. Since 2015, Eurostat annually assesses the level of digital skills of the EU population. Statistics Netherlands (CBS) provides an overview of the results for the Netherlands, enriched with several demographic characteristics (Statistics Netherlands, 2020). This overview shows that, on average, digital skills are higher for males, younger persons and higher educated persons. There also is a correlation between income and digital skills as those in the highest 20 percent income group tend to have better digital skills. However, Eurostat only publishes self-reported digital skills, which is prone to measurement error. Most of the extant scientific literature also uses self-reported skills. A notable exception is the work by Van Deursen & Van Dijk (2011). Van Deursen and Van Dijk use computer assignments to measure digital skills and show that these are negatively related with age and positively with education. However, this study excluded persons with no or very limited computer experience.

Thus far, the scientific literature on digitalization and the labor market has mainly focused on implications for employment and wages at an aggregate level. The extensive literature on job market polarization shows an increase over time in employment for low-pay and high-pay jobs and a decrease in employment for middle-pay jobs, in both the U.S. and in Europe (see e.g. Autor et al. (2006), Goos & Manning (2007) and Goos et al. (2009)). At the same time, wage inequality is rising and especially middle wages are staying behind (see e.g. Acemoglu & Autor (2011) and Van Reenen (2011) for international evidence and Groot & De Groot (2011) for an analysis of the Netherlands). Many authors point to routine biased technological change as an explanation for the worsening position of middle-pay workers (see e.g. Autor et al. (2003), Michaels et al. (2013), Goos et al. (2014), Autor et al. (2008)). According to this theory, ICT substitutes for routine cognitive tasks (e.g. clerical and administrative occupations) that are typically middle-paid, but complements high-skilled cognitive non-routine tasks (e.g. teachers, doctors) and has only little impact on non-routine manual workers (e.g. hairdressers, janitors).

The topic of labor market returns to digital skills relates to the broader topic of economic returns to human capital. Hanushek & Woessmann (2008) provide a discussion of the literature. Following Mincer (1970), many studies focus on the returns to schooling as this variable is readily available in many datasets. However, as Hanushek & Woessmann (2008) argue, schooling is but one of the factors influencing human capital formation and the quality of schooling differs over time and between regions. Standardized cognitive tests, like the one used in PIAAC, seem a better proxy for human capital and have consistently been shown to be positively correlated with income even after controlling for schooling (see e.g. Hanushek et al. (2015) for recent evidence). Aside from cognitive skills, several authors have found that social skills are increasingly rewarded on the labor market and increasingly complement cognitive skills (see e.g. Deming (2017) and Weinberger (2014)). This is probably related to digitalization because computers can take over certain cognitive tasks, but they generally perform poorly on tasks that require social skills.

3 Data

We use the Dutch part of the OECD Survey of Adult Skills from the Programme for the International Assessment of Adult Competencies (from now on referred to as PIAAC data) in combination with administrative data from Statistics Netherlands (from now on referred to as CBS data). Statistics Netherlands provided us access to the academic use PIAAC dataset. This dataset contains a unique person identifier that allows us to link the PIAAC data to administrative labor market data from CBS.

The PIAAC data provide measures of the skills of the respondents. Additionally, from the PIAAC data we take several background characteristics and the labor market situation at the time of the PIAAC survey. From the CBS data we take more detailed labor market statistics in the years after the PIAAC survey, which will be our main dependent variables in our analysis of labor market outcomes. Table 8 in the Appendix lists all variables we use and their source.

3.1 PIAAC data

In the Netherlands, the OECD gathered data between September 2011 and March 2012 (OECD, 2013). The survey administrators invited 10,255 persons between 16 and 65 years old who lived in the Netherlands at that time to participate in the survey. Of the invited persons, 5,170 took part in the survey. Interviewers visited each participant on his/her home address to take the questionnaire.

The PIAAC survey consists of multiple parts. First, the participant answers background questions on demographics, education and labor market situation. Next, the interviewer asks if the respondent has any computer experience and whether (s)he is willing to take a skill test on a computer. In case of no computer experience or refusal of a computer-based test, the interviewer provides the respondent with a paper test of literacy and numeracy, skipping digital skills. The remaining respondents first take a basic ICT skill test. If the respondent fails this test, (s)he also takes the paper test. Figure 8 in the Appendix visualizes the flow of the questionnaire.

The skill test measures the respondents' skills in the domains of literacy, numeracy and problem solving in technology-rich environments. To measure the last skill, respondents perform several tasks using mock-up websites and spreadsheets. This test aims to measure the cognitive skills required in the information age, such as deciding which information is needed, critically evaluating this information and using it to solve problems. In the remainder of this paper, we refer to problem solving in technology-rich environments as digital skills. The PIAAC survey uses a complex questionnaire design for skill measurement, resulting in so-called plausible values for the skills. We give a concise overview in the Appendix, please refer to OECD (2016) for more details.

3.2 CBS data

The CBS data consist of two separate administrative datasets.[2] The first contains monthly data on job characteristics and wages of all salaried persons who are employed in Dutch companies and covers January 2012 till September 2019.[3] From this dataset we derive the yearly average gross hourly wage and the contract type (fixed or temporary) in December of each year.[4]

The second dataset from Statistics Netherlands contains monthly socio-economic data for the entire Dutch population from January 2012 till December 2018. The data categorize persons based on their main source of income, for example employee, self-employed, recipient of unemployment benefits, recipient of social welfare benefits and without income. Using these data, we define a person as employed when (s)he is employee, self-employed, director with majority shareholding, or otherwise employed. We construct a dummy variable that indicates the employment status (employed or not employed) in December of each year. We also construct a dummy variable that indicates whether a person has been active on the labor market during 2012-2018. We define an individual to be active when he or she has been employed for at least one month during 2012-2018.

Lastly, using the socio-economic data from Statistics Netherlands, we construct a dummy variable that indicates whether a person has had at least one unemployment spell in 2012-2018. Here we define someone as unemployed when (s)he receives unemployment benefits or social welfare benefits. Recipients of those benefits have a legal obligation to seek work. Other non-working persons, e.g. homemakers or retired persons, are in general not applying for jobs, making the concept of an unemployment spell irrelevant for them.

[2]Statistics Netherlands refers to the datasets we use as Spolisbus and Secmbus.

[3]We accessed the data in February 2020. At that moment, data on the last quarter of 2019 were not yet available.

[4]For 2019 we take the average wage in January-September and the contract type in September.

3.3 Final dataset

From the 5,170 participants in the PIAAC questionnaire, we remove 87 persons with literacy-related non-response and we remove six persons with either missings in the relevant background variables or who could not be traced in the socio-economic CBS data. This gives a final PIAAC dataset with 5,077 persons. Subsequently, we link those individuals to the corresponding CBS data.

The socio-economic data can be fully matched for January 2012, but over time some persons emigrate or pass away, leading to a decline in matches over time. For December 2018, we can still match 4,939 of the PIAAC participants to the socio-economic data.

For the wage data we have less matches as this administrative dataset only contains salaried persons. The number of matches declines over time, from 3,800 in 2012 to 3,033 in 2018, but increases slightly in 2019. Note that the persons who were matched in 2018 were not necessarily matched in 2012 as well, because some, for example, might have been jobless or self-employed in 2012 and have a salaried job in 2018.

4 Methodology

We start our analysis with a description of the full sample, broken down by digital skill level. This shows the demographic differences between persons with varying levels of digital skills. Next, we analyze the relation between digital skills and labor market position. In this analysis we follow e.g. Hanushek et al. (2015) and focus on prime-age persons who were between 30 and 55 years old in 2012.

The PIAAC data contain two groups of persons with missing digital skills. The first group consists of 222 persons who prefer a paper-based assessment. In our main labor market analysis, we leave those people out, because we do not know the reason why they prefer a paper test. The second group encompasses 311 persons who failed the basic ICT test. For those persons we impute a low digital skill. The skills are measured in a range from zero to 500. The imputed skills are random draws from a Uniform [120, 220] distribution, which is approximately the range of the bottom five percent of digital skills.[5] As sensitivity checks, we will also estimate the labor market models including the persons who prefer a paper-based test, excluding the persons who failed the basic ICT test or including all age groups.

For the labor market analysis, we estimate a Mincer-type model

$$Y_i = \alpha + \beta_1 Digi_i + \beta_2 Num_i + \beta_3 Lit_i + \gamma X_i + \epsilon_i \tag{1}$$

where Y_i is the outcome (e.g. log hourly wage, employment) of individual i, $Digi_i$, Num_i and Lit_i signify the digital, numeracy and literacy skills, respectively, and X_i are background variables. We normalize the skill variables to have a mean of zero and standard deviation of one, such that β_1, β_2 and β_3 measure the effect of a one standard deviation change in skills. As background variables we use the age in 2012, education level, whether the respondent is born in the Netherlands, whether the respondent has children, gender and an interaction term between children and gender. We add this interaction between gender and children because it is common in the Netherlands for females with children to not work fulltime (see e.g. Bosch et al. (2010)). We add both literacy and numeracy and education level to the model to reduce omitted variable bias, but explore other model specifications as a sensitivity check.

[5] We tried other distributions and ranges, but this did not change our main results.

It is important to keep in mind that the outcome variables that originate from the CBS data are available for multiple years while the skills and background variables are only measured in 2012. We estimate equation (1) separately for each year. In our main analysis, we use a linear model for each outcome variable, but as a sensitivity check we will estimate a logit model for the binary outcomes (e.g. employment).[6] As the skills are measured in plausible values, we use the Repest package in Stata to estimate Equation (1). This package gives unbiased standard errors, whereas a standard OLS estimation would generally underestimate the standard errors. We provide more details in the Appendix.

[6]As the skills and background variables are measured at a single point in time, we cannot estimate a fixed-effects panel model. A random-effects panel model cannot be estimated either because the standard errors need a correction for the fact that the skills are measured in plausible values and the available software cannot combine this with random effects.

5 Data description

5.1 Description full sample

In Table 1 we show descriptives for different subgroups in the sample, based on their digital skills. In this table, we use the PIAAC skill classification of four levels for digital skills, see Chapter 21 of OECD (2016) for more details. In this classification, level 1 can be interpreted as basic skills, that is, the ability to use widely available and familiar applications like email, with little or no navigation required to solve the task at hand. Persons classified into level 0 lack even these basic skills. Persons assigned to levels 2 and 3 are able to solve tasks that require the use of more specific applications and tools (e.g. the sort function in a spreadsheet) and that involve multiple steps. About one third of the sample has basic skills (level 1) and another third has above basic skills (level 2). Almost eight percent of the sample is highly skilled (level 3), while the remaining 24 percent lacks even basic skills or prefers a paper test.

Table 1: Descriptives by digital skill level, full sample

Digital skill level	(1) **Prefer paper test**	(2) **Failed basic test**	(3) **Level 0**	(4) **Level 1**	(5) **Level 2**	(6) **Level 3**	(7) **Full sample**
Frequency							
% of sample	4.58	6.82	12.33	32.95	35.47	7.84	100
Background variables							
Age							
Average age	49.83	49.07	48.13	42.22	36.57	32.60	41.72
% 30+	91.14	87.98	87.68	76.99	65.97	56.53	75.02
% 40+	82.30	78.43	75.01	59.83	42.21	26.53	57.85
Gender							
% Male	41.65	47.81	42.84	48.46	53.56	60.81	49.20
% Female	58.35	52.19	57.16	51.54	46.44	39.19	50.80
Education[a]							
% Less than high school	54.83	69.80	52.64	33.39	16.57	7.26	31.18
% High school	31.71	21.73	34.27	42.83	40.45	34.54	38.44
% Above high school	13.45	8.47	13.09	23.78	42.98	58.20	30.38
Children							
% No children	24.37	28.84	22.69	34.37	48.63	62.29	36.97
% Yes	75.63	71.16	77.31	65.63	51.37	37.71	63.03
Born in the Netherlands							
% Yes	70.82	63.54	79.75	89.35	92.88	93.18	90.90
% No	29.18	36.46	20.25	10.65	7.11	6.82	9.10
Skills							
Average numeracy score[b]	248.30	214.12	227.34	272.91	309.32	340.82	280.40
Average literacy score[b]	256.62	226.63	226.60	275.26	313.17	345.50	284.05

[a] The education category 'Foreign qualification' has been left out for privacy reasons as this category is very small.
[b] Numeracy and literacy are measured on a scale from zero to 500.

Not surprisingly, persons with low digital skills are on average older. However, about 23 percent of the persons who failed the basic test or have below basic skills are below 40 years old and about 12 percent of those low-skilled persons even are below 30 years old. This implies that although in a few years' time many of the people who lack basic digital skills will retire from the workforce, a non-negligible part of the low-skilled still has 25 or more productive years to go.

Persons with higher digital skills are more often male, on average have a higher education and are more often born in The Netherlands. They are also less likely to have children, which might be related to their lower average age. There is a clear positive correlation between digital skills and literacy and numeracy. The ability to navigate through a digital application obviously requires a certain level of literacy and numeracy skills. Also, the PIAAC measurement of digital skills focuses on cognitive skills like information processing in a digital environment. Although this is not the same as literacy and numeracy, the different skills are clearly related.[7]

Finally, we note that the demographic characteristics of the group preferring a paper test are similar to the characteristics of the groups who failed the basic test or do not have basic skills. This suggests that the majority of those who choose a paper test do so because they are not familiar with computers. Surprisingly, the average numeracy and literacy scores of those who prefer a paper test are relatively high. This points to a group of persons with average cognitive skills who never had the opportunity to familiarize themselves with computers.

5.2 Description sample for labor market analysis

In Table 9 in the Appendix we show the frequency distributions of digital, numeracy and literacy skills in the different samples that we use in our analysis of the labor market. We again use the PIAAC skill classification into four levels for digital skills and we use the PIAAC classification into six levels for literacy and numeracy (OECD, 2016). As noted before, the numeracy, literacy and digital skills are correlated. In our main sample, the correlation between numeracy and literacy skills is 0.88, the correlation between numeracy and digital skills is 0.75 and the correlation between literacy and digital skills

[7]We also estimated an ordered logit model with the digital skill groups as outcome variable and the background variables and literacy and numeracy scores as explanatory variables. All variables have the expected sign. Literacy, numeracy, age and education are highly significant. Gender, children and born in The Netherlands are not significant. Persons born outside The Netherlands on average have lower literacy and numeracy skills, which might explain the non-significance of this variable. Similarly, those with children on average are older. Regression outcomes are available on request.

is 0.78. Although the correlation is high, VIF values for our regressions do not indicate major multicollinearity issues.[8] Moreover, we do find significant results, while in general multicollinearity leads to inflated standard errors (see e.g. Greene (2018)).

In Table 10 in the Appendix we provide summary statistics of the background variables in the different samples we use. Note that excluding the respondents who failed the basic ICT test (column (3)) or including the respondents who prefer a paper-based test (column (4)) hardly changes the summary statistics as each of those groups only makes up about five percent of the prime-age sample.

[8] We find VIF values between three and 5.5. A general rule of thumb is that there are multicollinearity issues if the VIF value is above ten.

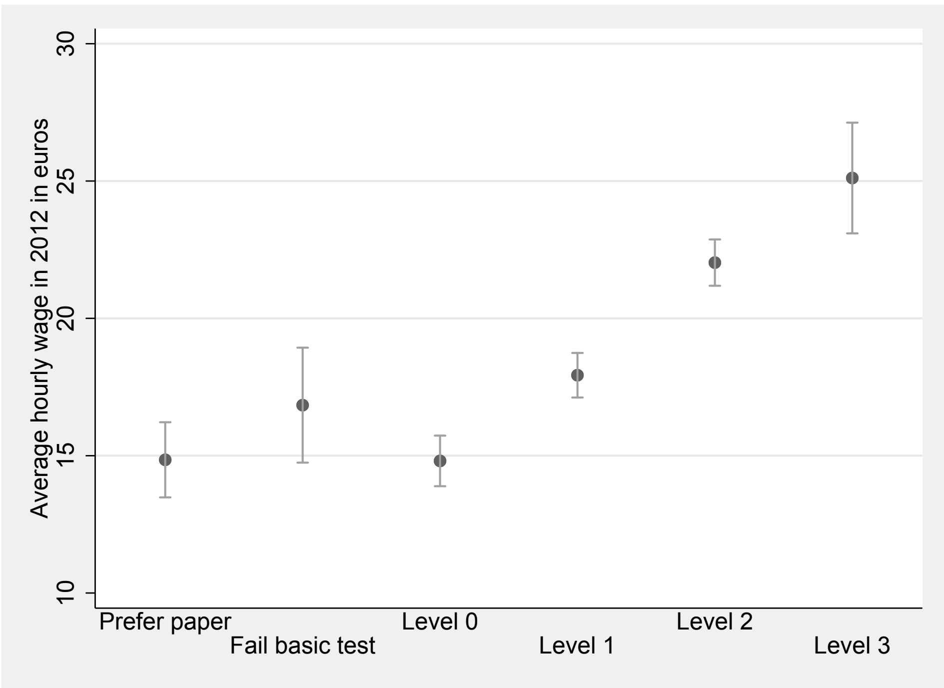
6 Digital skills and wage

Figure 1 shows the average hourly wage in 2012 for different levels of digital skills, for respondents between 30 and 55 years old in a salaried job. Although the figure does not correct for background characteristics, it does suggest that the groups that prefer a paper-based test or have a zero digital skill level have a similar wage level. The average wage of those who failed the basic ICT test seems slightly higher, but the confidence interval is relatively large and at 95 percent level the wage is not significantly different from the wage of the groups that prefer a paper-based test or have a zero digital skill level.[9] For the groups with skill levels 0 till 3, the wage increases roughly linear with digital skills.

Figure 2 plots the development of the nominal hourly wage over the years 2012-2019 for different skill groups. For each group, the wage increases roughly linearly, except for

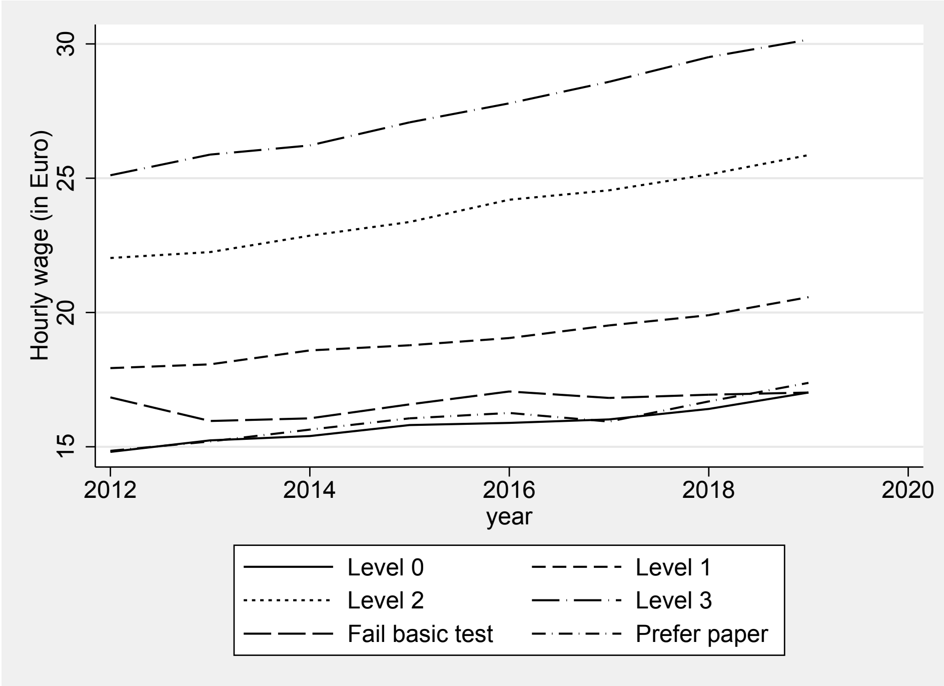
[9] An F-test of equivalence of the wage of the 'Fail basic test' group versus the 'Level 0' group has a p-value of 0.0822. An F-test of equivalence of the wage of the the 'Fail basic test' group versus the 'Prefer paper' group has a p-value of 0.1195. An F-test of equivalence of the wage of all three groups together has a p-value of 0.2076.

Figure 1: Hourly wage in 2012 by digital skill category (95 percent confidence intervals)



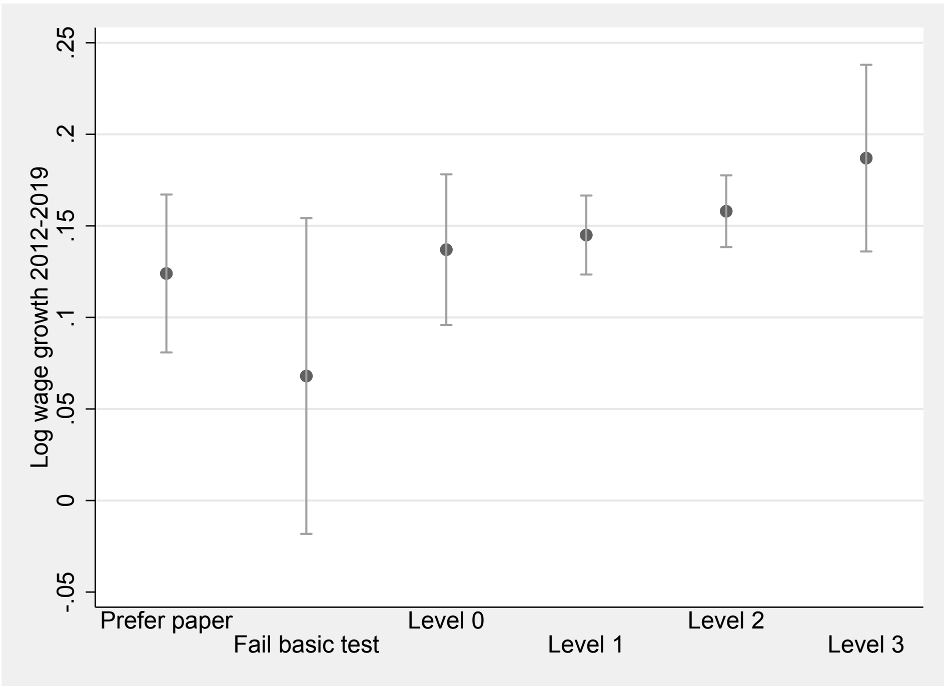
Source: Statistics Netherlands (CBS) and PIAAC

Figure 2: Nominal hourly wage over time by digital skill category



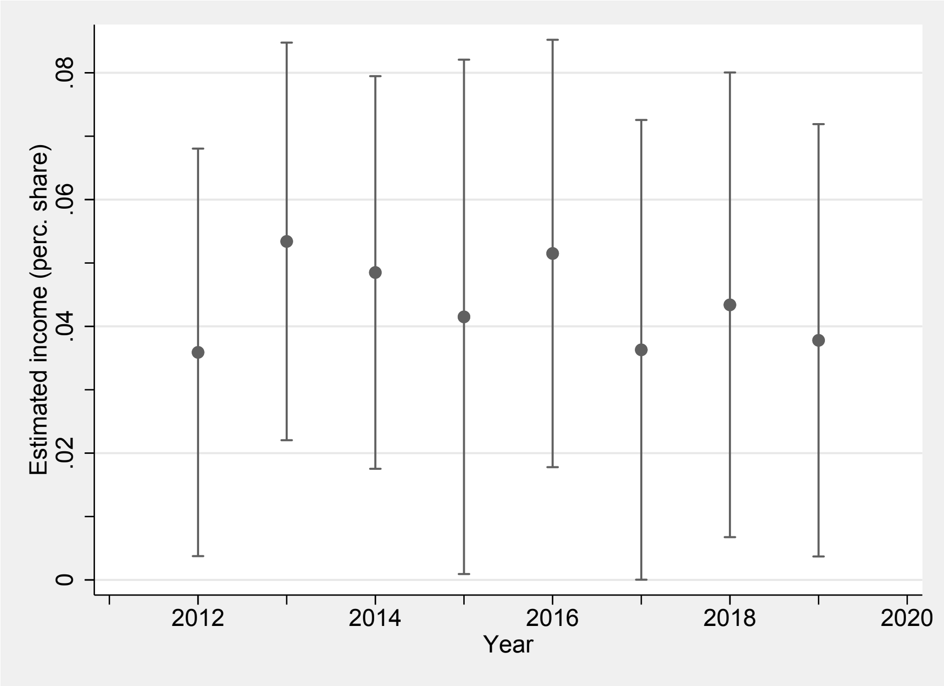
Source: Statistics Netherlands (CBS) and PIAAC

Figure 3: Log wage growth by digital skill category (95 percent confidence intervals)



Source: Statistics Netherlands (CBS) and PIAAC

Figure 4: Estimated coefficients digital skills with 95 percent confidence intervals (dependent variable: log hourly wage)



the group that failed the basic ICT test. To gain more insight into the wage growth, Figure 3 gives the average (log) wage growth between 2012 and 2019 for the different skill groups. The different groups have very similar growth rates. An F-test shows no statistically significant difference between all groups jointly (p=0.15).[10]

We estimate Equation 1 with as dependent variables the log hourly wage in the years 2012-2019 and the log wage growth. The full results are reported in Table 11 in the Appendix. Figure 4 shows the coefficients of digital skills for the log hourly wage in the years 2012-2019, with 95 percent confidence intervals. A one standard deviation increase in digital skills is related to an increase in the hourly wage of roughly four percent. All estimates are significant at 95 percent level. The estimates fluctuate slightly over time, but not significantly and not in a systematic pattern. The regression on the wage growth

[10]Note that the confidence interval for the group that failed the basic test is very large. When we winsorize the data, the log wage growth of those who failed the basic test is very similar to the log wage growth of those who prefer the paper test. Results of this exercise are available upon request. We also performed F-tests on equivalence of the wage growth of the group that failed the basic test versus the groups that prefer a paper test or have level zero. These tests show no statistically significant differences (p=0.67 and p=0.16 respectively).

confirms this. The coefficient on digital skills in this regression is very small and not significantly different from zero.[11]

The coefficients of numeracy and literacy in general are not significantly different from zero. For the background characteristics, we find that the log of age is highly significant, implying that the hourly wage increases stronger in age for young persons.[12] Note that the effect of age decreases over time, as the population ages. Indeed, the log of age is negatively and significantly related to wage growth. Education is strongly positively related to wage, and those not born in The Netherlands earn less. As expected, females with children earn significantly less than males with children. However, there is no significant difference in the wages of childless males and females. Wage growth is not related to any of those background characteristics, except that those with a foreign qualification tend to have a stronger wage growth. It might be that work experience built up over time helps them to prove the value of their qualification.

In the analysis above, we imputed low digital skills for persons who failed the basic ICT test. The imputed values are on average slightly lower than the skills in level 0, which theoretically seems a reasonable choice. However, Figure 2 suggests that the wage of persons who failed the basic ICT test is slightly higher than the wage of persons with skill level 0. As a sensitivity check, we estimate the models for log hourly wage in 2012-2019 including an additional dummy for the group who failed the basic ICT test. Figure 5 shows that this increases the coefficient on digital skills. Previously, a one standard deviation increase in digital skills was related to an increase in the hourly wage of roughly four percent. Now this is about six percent. For all years, the additional dummy is positive and for four out of the eight years, the dummy is significant at five percent level. This suggests that the persons who failed the basic ICT test earn a higher hourly wage than their imputed skills might suggest.

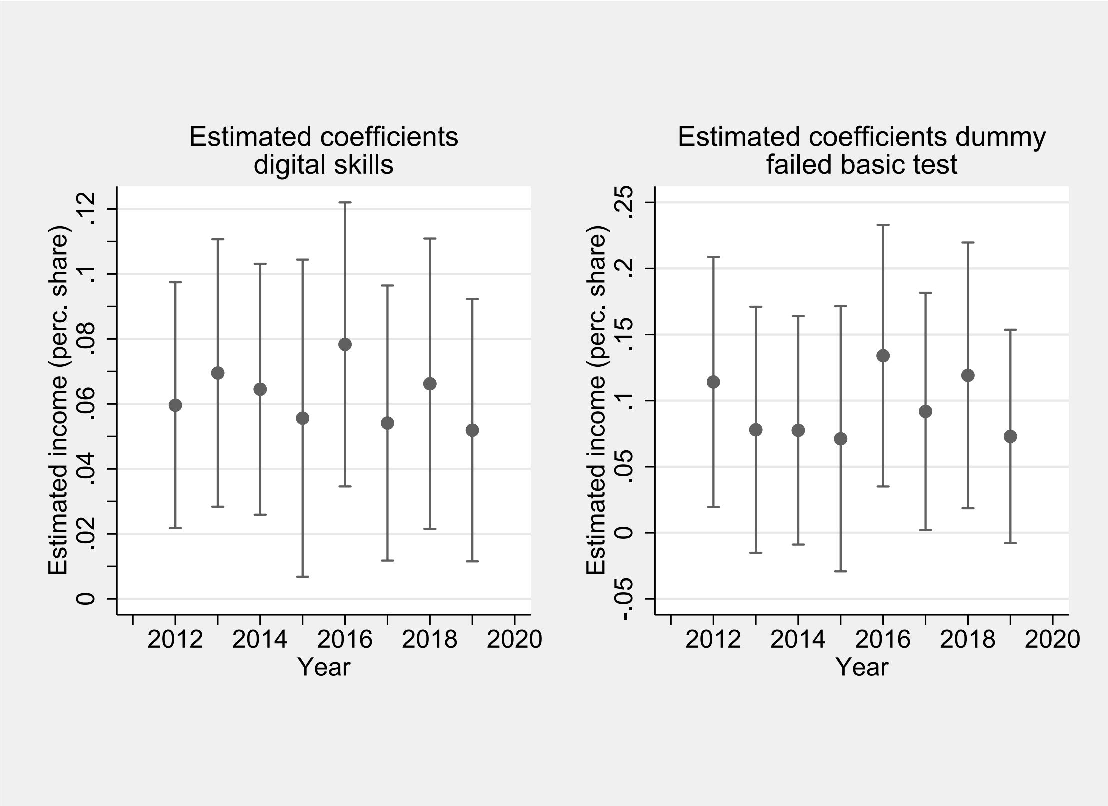
Another way to test this, is to estimate the models on the log hourly wage in 2012-2019 including dummies for all digital skill groups, instead of the linear continuous specification used before. Table 2 gives the estimated coefficients for those dummies.[13] The group that failed the basic ICT test is the reference category. In seven out of eight years, the group with skill level 0 has a negative estimate, again suggesting that this

[11] We also analyzed other operationalizations of wage growth, such as the average year-on-year growth, but this gave similar results.

[12] We also estimated the models using a linear or quadratic specification of age or with dummies for five-year age brackets. Those specifications gave a worse fit.

[13] Full results are available on request.

Figure 5: Estimated coefficients digital skills and dummy for those who failed the basic test, with 95 percent confidence intervals (dependent variable: log hourly wage)



group has a lower log hourly wage than the group that failed the basic test. However, the difference between those groups is never statistically significant. The coefficients for the groups with skill levels 1, 2 and 3 indicate a roughly linear increase in log hourly wage with the digital skill levels. In general, the coefficients are not significant. However, when we take skill level 0 as reference category, levels 2 and 3 are significant for most years.

In the Appendix, we discuss several other sensitivity checks where we either use a different sample or change the model specification. The results in general are robust to those sensitivity checks.

Table 2: Sensitivity check wage model: nonlinear digital skills

	2012		2013		2014		2015	
	coeff.	se	coeff.	se	coeff.	se	coeff.	se
Digital skills: level 0	-0.0565	(0.0470)	-0.0055	(0.0453)	-0.0147	(0.0401)	0.0030	(0.0449)
Digital skills: level 1	-0.0169	(0.0425)	0.0349	(0.0364)	0.0284	(0.0371)	0.0085	(0.0454)
Digital skills: level 2	0.0430	(0.0492)	0.1033**	(0.0425)	0.0911**	(0.0429)	0.0716	(0.0541)
Digital skills: level 3	0.0746	(0.0653)	0.1508**	(0.0637)	0.1202*	(0.0662)	0.1002	(0.0784)
Observations	2,120		2,088		2,060		2,038	

	2016		2017		2018		2019	
	coeff.	se	coeff.	se	coeff.	se	coeff.	se
Digital skills: level 0	-0.0570	(0.0498)	-0.0431	(0.0464)	-0.0543	(0.0482)	-0.0097	(0.0373)
Digital skills: level 1	-0.0056	(0.0372)	-0.0019	(0.0417)	-0.0092	(0.0434)	0.0051	(0.0380)
Digital skills: level 2	0.0645	(0.0444)	0.0435	(0.0488)	0.0407	(0.0504)	0.0539	(0.0458)
Digital skills: level 3	0.0956	(0.0718)	0.0718	(0.0683)	0.0823	(0.0668)	0.0949	(0.0650)
Observations	2,003		1,912		1,801		1,971	

Notes: Standard errors in parentheses. Reference category: individuals who failed basic ICT test.
All other covariates from main specification included.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

7 Digital skills, employment and labor force participation

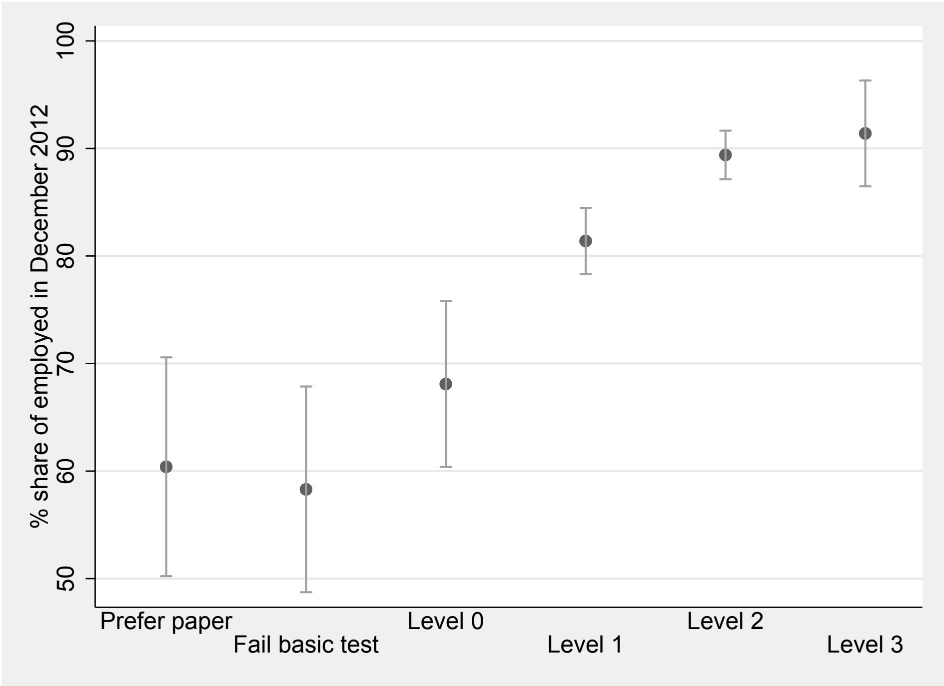
7.1 Employment

Figure 6 shows the employment share in December 2012 for different levels of digital skills for respondents between 30 and 55 years. The employment share in general increases with skill level. However, the groups with skill levels 2 and 3 do not differ much as both have an employment share of around 90 percent. Moreover, the groups that failed the basic ICT test or prefer a paper-based test also have a similar employment share, around 58 percent. Figure 7 shows the development of the employment share over time by digital skill category. Over time, the employment share is fairly stable.

We estimate Equation 1 with a dummy for employment in December as dependent variable for the years 2012-2018. The full results are in Table 12 in the Appendix. For easy reference, Table 3 shows the estimated coefficients for digital skills. As Figures 6 and 7 show that the relation between digital skills and employment is likely nonlinear, we include dummies for the digital skill levels instead of a continuous variable. We use the group who failed the basic ICT test as reference category.

As expected, the coefficients increase with skill level, except for level 3. In four out of seven years, the coefficient for level 2 is significant at five percent level and in the remaining years the coefficient is significant at ten percent level. In the first years, persons with digital skills level 1 have a roughly eight percent higher employment share than persons who failed the basic ICT test. For digital skills level 2 this is about ten

Figure 6: Percentage share of employed in December 2012 by digital skill category (95 percent confidence intervals)



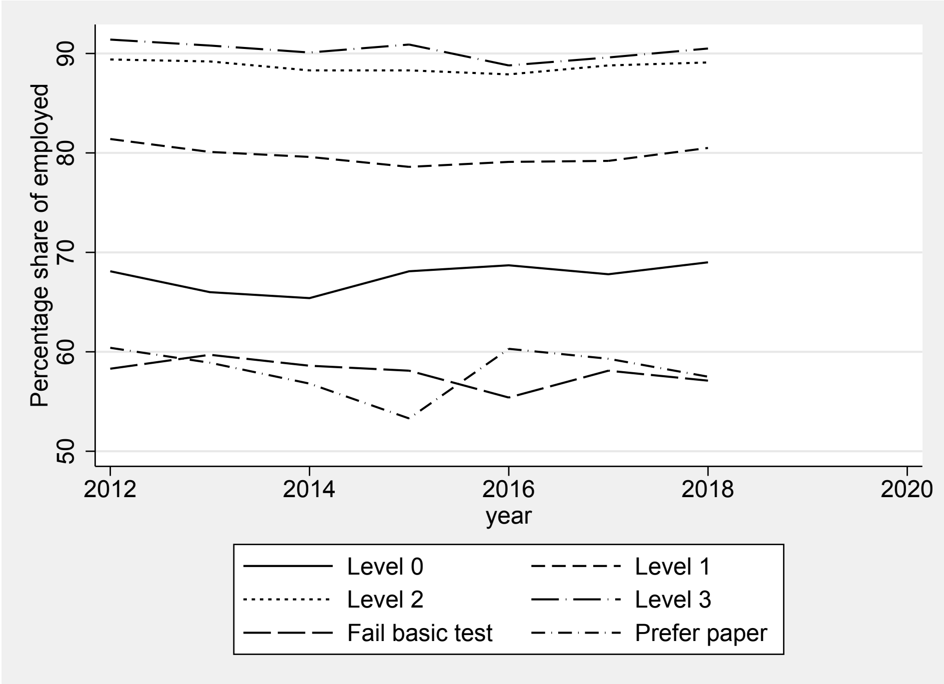
Source: Statistics Netherlands (CBS) and PIAAC

Table 3: Baseline employment regressions 2012-2018

Dependent variable Employed at the end of year	(1) **2012**	(2) **2013**	(3) **2014**	(4) **2015**	(5) **2016**	(6) **2017**	(7) **2018**
Digital skills: level 0	0.0447	0.0143	0.0174	0.0565	0.0858	0.0541	0.0851
	(0.0606)	(0.0591)	(0.0589)	(0.0587)	(0.0629)	(0.0614)	(0.0591)
Digital skills: level 1	0.0966*	0.0737	0.0833	0.0742	0.1130**	0.1087**	0.1405***
	(0.0511)	(0.0494)	(0.0518)	(0.0501)	(0.0521)	(0.0510)	(0.0500)
Digital skills: level 2	0.1115**	0.0987*	0.1058*	0.1004*	0.1367**	0.1469**	0.1689***
	(0.0551)	(0.0580)	(0.0574)	(0.0524)	(0.0591)	(0.0575)	(0.0553)
Digital skills: level 3	0.0869	0.0711	0.0835	0.0814	0.1002	0.1143*	0.1400**
	(0.0701)	(0.0763)	(0.0763)	(0.0698)	(0.0692)	(0.0683)	(0.0647)
Observations	2,608	2,602	2,595	2,588	2,582	2,576	2,563

Notes: Standard errors in parentheses. Reference categories: digital skills- failed basic ICT test. Full estimation results are reported in the appendix.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 7: Employment share over time by digital skill category



Source: Statistics Netherlands (CBS) and PIAAC

percent.[14] In later years, the coefficients for digital skills are larger, but the increase in the coefficients is limited to roughly one standard error and, hence, is not statistically significant. In most years, the coefficient for digital skills level 3 is not significant. Not only is the coefficient slightly smaller than for level 2, the standard error also is higher as the group with level 3 is relatively small.

The numeracy and literacy skills in general are not significant. Furthermore, we find that employment increases with education level and is lower for persons who were not born in The Netherlands. In the first few years, there are no significant age effects, but from 2015 onward the oldest age group starts to drop out of employment. Note that the group with 51-55 years of age in 2012 is 56-60 years old in 2017 and some of those people probably enter early retirement. In the first few years, females with children work less often, but over time this coefficient becomes less significant. At the same time, the coefficient on females becomes negative and significant. One explanation for this might be the fact that our variable on having children is based on the situation

[14]Note that these effects are much smaller than in the raw data as presented in figure 7. The background variables explain part of the difference between the skill groups.

in 2012. It is likely that in 2017-2018 more women in our sample have given birth and stopped working, something that will not be picked up by the 'children' variable, but will be reflected in the coefficient on gender. An additional explanation might be that the coefficient on the interaction effect is likely driven by older cohorts. In older generations, childbirth more often was a reason to fully stop working, while younger generations tend to switch to part-time work.[15] Over time, also the males in the older generation drop out of the labor force, thus weakening the interaction term.

In the Appendix, we discuss the results of a range of sensitivity checks on different samples, different model specifications and alternative employment measures.

7.2 Labor force participation

The results on employment indicate that individuals with below-basic skills are significantly less often employed than those with above-basic skills. However, the results do not state the reason why employment is lower. It might be that those with below-basic skills are more often on temporary jobs with unemployment spells in between, but it might also be that they are less often in the labor force. The administrative data from CBS does not include labor force participation, as this is usually measured using questionnaires. Our dummy variable *active*, which indicates whether someone has been employed (either in a salaried job or in self-employment) for at least one month in the period 2012-2018, seems a reasonable approximation for labor force participation.[16] Of the respondents between 30 and 55 years, about half of those who were unemployed in December 2012 are not active. For other years, this fraction is similar.

We estimate Equation 1 with the dummy variable *active* as dependent variable. The full results are reported in column (1) in Table 13 in the Appendix. For easy reference, Table 4 gives the coefficients for digital skills. The results show that the group that failed the basic ICT test is significantly less likely to have been active on the labor market during 2012-2018 than the groups with skill level 1 and 2. The difference is about twelve percentage points. The difference with the group with skill level 0 is smaller, with eight percentage points, and is only significant at ten percent level. The difference with the group with skill level 3 is also smaller and not significant. This coefficient has a relatively high standard error due to the small size of the group with skill level 3.

[15] Consistent with this hypothesis, when we estimate the model on persons between 30 and 40 years old, the interaction term is small and non-significant even in 2012.

[16] As robustness check, we also did the analyses with labor force participation defined as being employed for at least 25 months in the period 2012-2018. The conclusions do not change when we use this definition.

The coefficients on the other explanatory variables show that the older age group (51-55 years old) is less likely to have been active on the labor market during 2012-2018. The same holds for females and for persons who were not born in the Netherlands. Persons with a high education level are more likely to have been active on the labor market during 2012-2018.

The results on the dummy variable *active* suggest that the low employment of the group that failed the basic test can at least partially be explained by a lower labor force participation. As an additional check, we re-estimate the models for employment in December, but now conditional on being active on the labor market at some point during 2012-2018. The results for employment in December 2012 are presented in column (2) in Tables 4 and 13, the results for other years are very similar and are available upon request. Restricting the sample to persons who have been active on the labor market at some point during 2012-2018 substantially decreases the size of the coefficients for digital skills. They are not significantly different from zero anymore. The coefficients of the other explanatory variables also decrease in size, but in general stay significantly different from zero. These results confirm that the previous results on employment are mostly due to a lower labor force participation of the group that failed the basic ICT test and not to more temporary jobs and unemployment spells in this group.

To test this in a more direct way, we estimate Equation 1 with unemployment spell as dependent variable. This variable indicates whether someone received unemployment benefits or social welfare during at least one month in 2012-2018. We estimate this model conditional on being active on the labor market for at least one month in 2012-2018.

Table 4: Regressions labor force participation, unemployment spells and contract type

	(1)	(2)	(3)	(4)
Dependent variable	**Active in 2012-2018**	**Employed in December 2012**	**Had unemployment spell**	**Fixed contract December 2012**
Conditional on active in 2012-2018	No	Yes	Yes	No
Digital skills: level 0	0.0861*	-0.0189	0.0678	-0.0533
	(0.0453)	(0.0617)	(0.0585)	(0.0747)
Digital skills: level 1	0.1220***	0.0027	0.0750	0.0113
	(0.0431)	(0.0488)	(0.0558)	(0.0633)
Digital skills: level 2	0.1219***	0.0142	0.0718	0.0414
	(0.0471)	(0.0495)	(0.0595)	(0.0686)
Digital skills: level 3	0.0900	0.0146	0.1062	0.0143
	(0.0558)	(0.0581)	(0.0733)	(0.0904)

Notes: Standard errors in parentheses. Reference categories: digital skills- failed basic ICT test.
*** p<0.01, ** p<0.05, * p<0.1

The results are in column (3) in Tables 4 and 13 and confirm our hypothesis: the group that failed the basic ICT test does not have significantly more unemployment spells than the other groups. In fact, the results suggest that the groups with above-basic skills have a *higher* probability of an unemployment spell, although this difference is not significant. The coefficients on the other explanatory variables show that the probability of an unemployment spell decreases with age, probably because younger persons more often are in a temporary contract. Higher educated persons and those with children have a lower probability of an unemployment spell and those not born in the Netherlands have a higher probability of an unemployment spell.

The results until now suggest that the group that failed the basic ICT test is not more often temporarily employed than the other skill groups. We test this in a direct way by estimating Equation 1 with *contract type* as dependent variable. This is a dummy variable that takes the value 0 if the contract is temporary and 1 if the contract is fixed. Column (4) in Tables 4 and 13 gives the estimation results for the contract type in December 2012. The results for other years are similar and are available on request. As expected, there is no significant relation between the level of digital skills and the contract type. As for the other explanatory variables, older people, people with children and persons born in the Netherlands more often have a fixed contract.

To conclude, we find that the relatively low employment of those with below-basic skills is mainly due to a lower labor force participation. Once we control for this effect, the relation between employment and digital skill category disappears. Moreover, we find no relation between digital skills and unemployment spells or the contract type, suggesting that those with below-basic skills are just as likely to have a stable job as those with above-basic skills.

8 Self-reported labor market outcomes

In this section we compare our baseline results for the wage and employment regressions as discussed in Sections 6 and 7 with the outcomes using wage and employment status as measured in the PIAAC questionnaire. The PIAAC questionnaire contains questions about hourly wage and whether the respondent had a paid job at the time the survey was taken. The upper part of Table 5 depicts the estimation results for the wage regression, and the lower part of the table shows the comparison for the employment regression. All comparisons are exclusively for 2012, as the PIAAC survey has only been conducted in the Netherlands in this year. The full estimation results for the wage and employment model are reported in the Appendix in Tables 14 and 15 respectively.

The first column of Table 5 gives the estimation results for all respondents with non-missing wage based on the PIAAC questionnaire and the second column shows the results for a sub-sample trimmed at the top and bottom first percentile of the wage

Table 5: Comparing baseline results with results on self-reported wage and employment as measured in PIAAC

Dependent variable log (hourly wage)	(1) **PIAAC data** **2012**	(2) **trimmed PIAAC data** **2012**	(3) **Register data (baseline)** **2012**
Digital skills	0.0554	0.0265	0.0359**
	(0.0397)	(0.0242)	(0.0164)
Observations	1,780	1,521	2,120

Dependent variable Employed (=1)	(4) PIAAC data **2012**	(5) Register data (baseline) **2012**
Digital skills: level 0	0.0563	0.0447
	(0.0559)	(0.0606)
Digital skills: level 1	0.1213**	0.0966*
	(0.0508)	(0.0511)
Digital skills: level 2	0.1424**	0.1115**
	(0.0559)	(0.0551)
Digital skills: level 3	0.1172*	0.0869
	(0.0693)	(0.0701)
Observations	2,608	2,608

Notes: Standard errors in parentheses. Income in column (1) is measured as gross hourly wage as reported by the PIAAC respondent. Employment as measured in PIAAC (column 4) is set to 1 if respondents indicated to have had a paid job at the time the survey was taken. The reference category for the employment model is people who failed the basic ICT test. Full estimation results are reported in the appendix.
*** p<0.01, ** p<0.05, * p<0.1

distribution.[17] In the third column, we include the baseline results from Section 6 which are based on CBS register data. Although we find a significant positive association between digital skills and hourly wage in the baseline model, the coefficient for digital skills in the PIAAC-based wage (trimmed) model is not statistically significant. One reason for this finding is the higher standard errors in the PIAAC-based model. Even when we consider the trimmed sample, the standard errors in column (2) are considerably larger than for the baseline specification. This observation is valid for all coefficients. A plausible explanation, also mentioned in (Gauly et al., 2019), is that the PIAAC-based wage is measured less precisely than the wage in the register data because the former is based on self-reported answers.[18]

The full results in Table 14 in the Appendix show several other differences between the baseline CBS-based results and the PIAAC-based wage model. When using the PIAAC wage measure, we find a significant positive relationship between literacy skills and wage. This finding is consistent with other studies who used a similar specification and the same wage measure (see e.g. Hanushek et al. (2015)), but differs from our baseline findings. The patterns for numeracy, age and education are similar to our baseline results. Interestingly, the negative association between female and wage becomes much stronger and statistically significant when we use the PIAAC-based wage measure. Those differences might perhaps be attributed to under-reporting of wages by female respondents and/or over-reporting by male respondents.

The lower panel of Table 5 shows similar patterns for the employment regression based on PIAAC employment as our main analysis in Section 7. Again, the coefficients increase in skill level except for the highest level, level 3. The group of highly skilled persons is relatively small which results in less efficient point estimates. The coefficients for the PIAAC-based employment measure in column (4) are slightly larger than the corresponding coefficients for the CBS-based employment measure, but not significantly so. Also regarding the other covariates, we find similar results for the PIAAC-based employment model.

This analysis demonstrates that caution is advised when using the wage measure from

[17] The sample size for the model using PIAAC wage does not equal the sample size when using CBS data. The PIAAC data suffers from non-response, whereas the CBS data excludes the self-employed. We decided to use all available data points. Our results, therefore, are driven by both non-response and measurement error. We leave the disentanglement of those effects for future research.

[18] Indeed, the standard deviation of the PIAAC-based wage is 0.74 for the full sample and 0.49 for the trimmed sample, whereas it is 0.41 for the CBS wage.

the PIAAC survey. This is consistent with the findings by Gauly et al. (2019), who warn of biased estimates due to measurement error when comparing wage measures from the PIAAC survey and (German) register data. As for employment, we find that the differences between the PIAAC-based measure of employment and the employment measures based on register data are rather negligible.

9 Discussion and conclusions

We study the distribution of digital skills in the Dutch population and the relation of these skills with labor market outcomes. Despite the Netherlands being one of the most digitalized EU countries, almost a quarter of our sample has below-basic skills or prefers a paper test over a computer test. As expected, these persons are on average older, lower educated and more often female. However, about 20 percent of this group is below 40 years old, and about ten percent even is below 30 years old. Hence, although low digital skills are more prevalent among older persons, there is a non-negligible group of persons with low digital skills who still have 25 or more years left on the labor market.

Our results on labor market outcomes sketch a mixed portrait. A one standard deviation increase in digital skills is associated with an increase in wages by four to six percent. However, we find no difference between those who failed a basic ICT test and those who passed this test but have below-basic skills. As for employment, we find that those with at least basic digital skills have a roughly ten percentage point higher employment rate than those who failed the basic test. This difference is mainly due to a lower labor force participation of those with no digital skills. We find no relation between digital skills and unemployment spells or contract type, suggesting no difference in job stability between those with below-basic skills and those with basic or above-basic skills.

To interpret these findings, it is important to keep in mind that we cannot claim causality. Both reverse causation and omitted variable bias might be present. In the wage regressions, these problems are likely limited. Hampf et al. (2017) explore causality in the case of PIAAC and conclude that simple least-squares estimates likely provide a lower bound of the true returns to skills. Indeed, our estimates are smaller than the estimates provided by the instrumental variable regressions in Falck et al. (2020). However, our results on labor force participation probably suffer from reverse causation. For individuals who work with a computer staying up-to-date is part of their job, while those outside the labor force have less opportunities to acquire and update their digital skills.

We do not find evidence of any divergence over time between the different digital skill groups. This indicates that the relationship between outcomes in 2012-2019 and the digital skills in 2012 is persistent. Based on the theory of routine-biased technological change, one might expect a divergence in wages and employment over time. However, we study a relatively short time period in a country where wage inequality historically has been relatively stable compared to other OECD countries (Groot & De Groot, 2011).

Moreover, we use digital skills as measured in 2012 and cannot control for changes in skills over time.

A unique aspect of our study is that we combine PIAAC questionnaire data with register data on labor market outcomes. This not only allows us to follow persons over time and to study labor force participation and job stability, but also allows for a comparison between PIAAC self-reported labor market outcomes and outcomes from register data. This comparison shows substantial differences between the estimation results based on the PIAAC self-reported wage and the results based on wages from register data. For employment, the regression results based on the PIAAC questionnaire are very similar to those based on the register data. These findings indicate that caution is needed when using self-reported wage measures.

Policymakers both in the EU and in the Netherlands have expressed concerns that persons with low digital skills are vulnerable both in the labor market and in society in general. Our research confirms that indeed there is a sizable group with a vulnerable position. Our research shows that in the Netherlands about a quarter of the persons aged between 16 and 65 has below-basic digital skills. Those persons participate less often in the labor force, making them dependent on either a working partner or on state benefits. Those who do work generally have a lower wage, which in combination with lower skills renders those persons vulnerable for financial problems. However, those with low digital skills are not more likely to be unemployed or in a temporary contract. Therefore, concerns that this group is in an unstable job position do not seem well-founded.

Finding a solution for those problems is not an easy task. Offering training in digital skills is an obvious policy option, but it is unclear whether this will lead to the desired results. Given that those with low digital skills in general have lower literacy and numeracy scores, a training might only have modest effects on digital skills and might only have a long-lasting effect when the skills are regularly put to use. As we find no difference in wages between the lowest digital skill groups, a modest increase in skills is not likely to lead to higher wages or better jobs for those with no skills at the outset. Then again, increasing skills further to a basic level might lead to better labor market perspectives. Moreover, acquiring digital skills might stimulate persons who are currently out of the labor force to start participating in the labor market. As we cannot claim causality, we are hesitant in strongly advising extensive digital skill training, but our results certainly give reason for (controlled) experiments and further research on the effect of skills training.

In the future, our research can be extended in several directions. Our results might differ between sectors, as the value of digital skills might differ dependent on how digitalized a sector is. Some of the literature on job market polarization distinguishes between sectors based on task content (e.g. non-routine manual work or routine cognitive work). This might be a starting point for further analysis. Also, the current COVID-19 situation has led to a sudden upsurge in digitalization as many companies turned to remote work to limit the spread of the virus. It is an open question whether this will increase the wage premium on digital skills.

References

Acemoglu, D. & Autor, D. (2011). Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics*, volume 4 (pp. 1043–1171). Elsevier.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The Polarization of the U.S. Labor Market. *The American Economic Review*, *96*(2), 189–194.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics*, *90*(2), 300–323.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333.

Avvisati, F. & Keslair, F. (2020). REPEST: Stata module to run estimations with weighted replicate samples and plausible values. `https://econpapers.repec.org/software/bocbocode/S457918.htm`.

Bosch, N., Deelen, A., & Euwals, R. (2010). Is Part-time Employment Here to Stay? Working Hours of Dutch Women over Successive Generations. *Labour*, *24*(1), 35–54.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

European Commission (2019a). The digital economy and society index (desi). `https://ec.europa.eu/digital-single-market/en/desi`. Accessed: 2020-05-13.

European Commission (2019b). European skills agenda. `https://ec.europa.eu/social/main.jsp?catId=1223`. Accessed: 2020-07-10.

European Commission (2019c). New commission report shows the importance of digital resilience in times of crisis. `https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1025`. Accessed: 2020-07-10.

Eurostat (2019). Self-employed persons. `https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20190430-1`. Accessed: 2020-07-16.

Falck, O., Heimisch-Roecker, A., & Wiederhold, S. (2020). Returns to ICT skills. *Research Policy*.

Gauly, B., Daikeler, J., Gummer, T., & Rammstedt, B. (2019). What's my wage again? Comparing survey and administrative data to validate earning measures. *International Journal of Social Research Methodology*, 1–14.

Goos, M. & Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics*, *89*(1), 118–133.

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *American economic review*, *99*(2), 58–63.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, *104*(8), 2509–26.

Greene, W. H. (2018). *Econometric Analysis*. Pearson.

Groot, S. & De Groot, H. (2011). Wage inequality in the Netherlands: evidence, trends and explanations. CPB Discussion Paper 186, CPB Netherlands Bureau for Economic Policy Analysis.

Hampf, F., Wiederhold, S., & Woessmann, L. (2017). Skills, earnings, and employment: exploring causality in the estimation of returns to skills. *Large-scale Assessments in Education*, *5*(1), 12.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*, 103–130.

Hanushek, E. A. & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, *46*(3), 607–668.

Lane, M. & Conlon, G. (2016). The impact of literacy, numeracy and computer skills on earnings and employment outcomes. *OECD Education Working Papers*, *No. 129*.

Levy, F. (2010). How Technology Changes Demands for Human Skills. *OECD Education Working Papers*, *45*.

Michaels, G., Natraj, A., & Van Reenen, J. (2013). Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years. *The Review of Economics and Statistics*, *96*(1), 60–77.

Mincer, J. (1970). The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach. *Journal of Economic Literature*, *8*(1), 1–26.

Ministry of Economic Affairs and Climate (2019). Nederlandse digitaliseringsstrategie 2.0 [dutch digitalisation strategy 2.0]. `https://www.rijksoverheid.nl/documenten/rapporten/2019/07/05/nederlandse-digitaliseringsstrategie-2.0`.

OECD (2013). Country note- netherlands: Survey of adult skills - first results. `http://www.oecd.org/skills/piaac/Country%20note%20-%20Netherlands.pdf`. Accessed: 2020-05-14.

OECD (2016). Technical report of the Survey of Adult Skills (PIAAC). `http://www.oecd.org/skills/piaac/PIAAC_Technical_Report_2nd_Edition_Full_Report.pdf`. Accessed: 2019-12-03.

Schel, M. & Vennekens, A. (2020). Digitale vaardigheden voor technologisch burgerschap [Digital skills for digital citizenship]. Policy report, Rathenau Institute.

Statistics Netherlands (2020). Digital skills and demographic characteristics. `https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83428NED/table?dl=3D8A0`. Accessed: 2020-07-20.

Van Deursen, A. & Van Dijk, J. (2011). Internet skills and the digital divide. *New media & society*, *13*(6), 893–911.

Van Reenen, J. (2011). Wage inequality, technology and trade: 21st century evidence. *Labour Economics*, *18*(6), 730–741.

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI monograph series*, *2*(1), 9–36.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, *96*(4), 849–861.

10 Appendix

10.1 Methodological details of the PIAAC skill measurement and subsequent statistics

The PIAAC data contain information on individual skill levels in a specific format, called plausible values. In this Appendix we describe the methodology behind plausible values and the statistics that are based on them. Please refer to OECD (2016) and Von Davier et al. (2009) for a more detailed explanation.

The PIAAC skill test is adaptive to prevent response fatigue. Respondents only answer a random selection of the questions and the level of difficulty of the questions is adapted during the test. This implies that the test scores of respondents cannot be directly compared, as different respondents got different questions. PIAAC uses an econometric model to estimate a posterior skill distribution for each respondent and each skill (literacy, numeracy and digital skills), based on the questions answered and background characteristics from the first part of the questionnaire.

For each respondent and skill, the PIAAC data contain ten random draws from this posterior distribution, the plausible values. The average plausible value indicates the skill level of the respondent. The spread in a respondent's plausible values indicates how precisely the skill was estimated. For example, when a respondent gave the wrong answer to some simple questions and a correct answer to more complicated questions, the spread in plausible values will be larger than when the simple questions were also correctly answered.

For all subsequent calculations (e.g. frequency distributions, regressions) that involve the skills, it is important to take into account that the skills are measured in the form of multiple random draws from a posterior distribution. While using only one plausible value in all calculations would give an unbiased result, the standard error would not correctly reflect the uncertainty in skill measurement. Therefore, all calculations should be based on all ten plausible values and should adjust the standard error based on the within person variation in plausible values.

For all our calculations, we use the Stata package Repest which was specifically designed to analyse data from PISA and PIAAC, see Avvisati & Keslair (2020). This package reports correct standard deviations taking into account that the skills are measured in the form of plausible values. Moreover, Repest also takes into account the sampling design of the PIAAC survey. In the Netherlands, the survey administrators employed a one-stage stratified design, using municipalities as stratum. The PIAAC data provides sampling weights to correct for nonresponse and provides replicate weights to adjust the standard errors for the stratified sampling design. The replicate weights are based on a Jackknife replication method, which gives heteroskedasticity-robust standard errors. The Repest package automatically incorporates the sampling weights and replicate weights. Below, we shortly explain how Repest handles plausible values, ignoring the sampling weights and replicate weights.

To correctly estimate equation (1) Repest estimates the model ten times, once for each draw of the plausible values. Next, Repest averages the resulting parameter estimates. Finally, Repest calculates the standard error of parameter θ with estimates $\hat{\theta}_m$ and corresponding standard errors $\hat{\sigma}_{\theta m}$ as

$$\sqrt{\frac{1}{M}\sum_{m=1}^{M}\hat{\sigma}^2_{\theta m} + \left(1+\frac{1}{M}\right)\sigma^2_W} \tag{2}$$

with M being the number of plausible values (ten in case of the PIAAC data) and σ^2_W the within variance of the ten estimates, that is,

$$\sigma^2_W = \frac{1}{M-1}\sum_{m=1}^{M}\left(\hat{\theta}_m - \frac{1}{M}\sum_{m=1}^{M}\hat{\theta}_m\right)^2. \tag{3}$$

To calculate the frequency distribution for the skills reported in Tables 1 and 9, we first recode each plausible value in a level. For example, for digital skills we recode a plausible value between 291 and 340 into level 2. Next, for each of the ten plausible values, Repest calculates the frequency distribution over the respondents. In the final step, Repest averages the ten frequency distributions, leading to the distributions as reported in Tables 1 and 9.

Note that the ten plausible values of a single person are not necessarily recoded into the same level. It might, for example, occur that a person has six plausible values in level 2 and four plausible values in level 3. Repest provides standard errors for the frequency distribution that reflect this uncertainty. For readability, we suppressed those in Tables 1 and 9 but they are available on request. Moreover, note that the frequency of those who failed the basic ICT test is higher than expected. There are in total 311 persons who failed the test. In the full sample, which contains in total 5,077 persons, this is 6.13 percent while Table 1 gives 6.82 percent. The reason for this discrepancy is that Repest automatically uses the PIAAC sample weights to correct the frequency distribution for nonresponse.

The calculations for the correlation between skills and for the average outcome (wage, employment, etc.) per skill level use a similar methodology as the one explained above. Repest calculates the correlation or average outcome for each of the ten plausible values and reports the average of the resulting ten correlations or average outcomes. Moreover, Repest adjusts the standard error in a similar way as in Equation (2) to reflect the uncertainty in the skill measurement.

10.2 Sensitivity checks for the wage regressions

In Table 6 we report the results of several sensitivity checks on the wage regressions. To save space, we only report the coefficient for digital skills. For easy reference, the first row of the results repeats the coefficients of the main model as discussed in Section 6.

First, we estimate a set of specification checks where we leave out the background variables, literacy and numeracy, or both. Not surprisingly, when we estimate the models only including the digital skills as explanatory variable, the coefficient is very large and highly significant. This coefficient also picks up the effects of literacy and numeracy and of the background variables such as age and education. When we include the background variables (but leave out literacy and numeracy), the coefficient of digital skills approximately halves, but still is about double the size of the coefficients in the main model. As mentioned before, digital skills, literacy and numeracy have a sizable positive correlation, which explains the higher coefficient for digital skills when literacy and numeracy are left out.[19] As a last specification check, we estimate the models including literacy and numeracy, but leaving out the background variables. For the first years, the coefficient for digital skills is very similar to the main regression, but for later years it becomes slightly larger. Interestingly, the coefficient for numeracy is large in this specification, suggesting that the background variables are stronger correlated with numeracy than with digital skills.[20] A (nonreported) check show that this is indeed true for the

[19] We also estimated models where we left out only literacy or only numeracy. In those models, the coefficient for digital skills is about 1.5 times higher than in the main regression. Results are available on request.

[20] The coefficient for numeracy is between 0.12 and 0.14 for the different years, compared to 0.03-0.05

variables gender and born in NL.

in the main specification.

Table 6: Sensitivity checks wage model

Type of check	Model		2012	2013	2014	2015	2016	2017	2018	2019
	Main regression (all skills + backgr. vars)	Coeff.	0.0359**	0.0534***	0.0485***	0.0415**	0.0515***	0.0363**	0.0434**	0.0378**
		se	(0.0164)	(0.0160)	(0.0158)	(0.0207)	(0.0172)	(0.0185)	(0.01879)	(0.0174)
Specification	Only digital skills	Coeff.	0.1640***	0.1723***	0.1760***	0.1804***	0.1899***	0.1940***	0.1982***	0.1911***
		se	(0.0109)	(0.0103)	(0.0109)	(0.012)	(0.0139)	(0.0119)	(0.0135)	(0.0102)
	Digital skills + backgr. vars	Coeff.	0.0820***	0.0916***	0.0912***	0.0909***	0.0977***	0.0937***	0.0970***	0.0844***
		se	(0.0119)	(0.0118)	(0.0125)	(0.0135)	(0.0146)	(0.0133)	(0.0144)	(0.0122)
	Digital skills + literacy and numeracy	Coeff.	0.0360*	0.0557***	0.0538***	0.0441**	0.0583***	0.0502***	0.0603***	0.0578***
		se	(0.0186)	(0.0167)	(0.0165)	(0.0214)	(0.0179)	(0.0186)	(0.0190)	(0.0184)
Sample	All ages	Coeff.	0.0380**	0.0517***	0.0309**	0.0278	0.0273*	0.0178	0.0224	0.0264*
		se	(0.0161)	(0.016)	(0.0151)	(0.0183)	(0.0156)	(0.0161)	(0.0162)	(0.015)
	No imputation	Coeff.	0.0622***	0.0771***	0.0630***	0.0489*	0.0775***	0.0547**	0.0668**	0.0458**
		se	(0.0227)	(0.0259)	(0.0234)	(0.0276)	(0.025)	(0.0242)	(0.026)	(0.0246)
	Full imputation	Coeff.	0.0353***	0.0445***	0.0415***	0.0349**	0.0428***	0.0356**	0.0419***	0.0344***
		se	(0.0135)	(0.0128)	(0.0137)	(0.0158)	(0.0132)	(0.0138)	(0.0137)	(0.0129)
	Winsorized income (top and bottom 1%-tile)	Coeff.	0.0393**	0.0460***	0.0506***	0.0456***	0.0436***	0.0282	0.0384**	0.0370**
		se	(0.0152)	(0.0133)	(0.0146)	(0.0175)	(0.0162)	(0.0175)	(0.0168)	(0.016)

Next, we perform a set of sample checks. In our main analysis in the text, we include all persons between 30 and 55 years old in our sample, but leave out those who prefer a paper test. When we include all ages, but still leave out those who prefer a paper test, the coefficient on digital skills decreases, especially for the later years. A further analysis (nonreported) shows that this is mainly due to the addition of young people to the sample. There are various explanations why the coefficient is lower for young people and decreases over time. First, the log of age in the background variables does not give a correct fit for the very low ages; the wage growth for persons till 25 years is stronger than the log of age can accommodate. When we use a more flexible specification with dummies for five-year age brackets, the coefficient for digital skills is consistently lower for the full age sample than for the age restricted sample, taking away the time effect. Second, wages do not strongly reflect skills for persons at the start of their career, see e.g. Hanushek et al. (2015), leading to lower coefficients. Third, the youngest persons in our sample might still further develop their digital skills, weakening the relation between skills and wage.

We return to the sample of 30-55 year old persons, but next leave out those who failed the basic ICT test ('no imputation'). This increases the coefficient on digital skills. As expected, the results are similar to the results when we add a dummy to the model for the group that failed the basic test (see Figure 5 in the main text). When we include both those who failed the basic test as well as those who prefer a paper test ('full imputation'), the coefficients on digital skills are fairly similar to those of the main model, albeit in some years slightly smaller.

As a final sensitivity check, we winsorize the wage by removing the top and bottom one percent. We note that outliers are rare in the register data that we use and that taking the log of the hourly wage as dependent variable seems sufficient to accommodate for the wage distribution. This is confirmed by the outcome of the regression on winsorized data. The coefficients on the digital skills are similar to the coefficients on the main regression and are not structurally lower or higher.

10.3 Sensitivity checks for the employment regressions

In Table 7 we report the results of a range of sensitivity checks on the employment regressions. For readability, we only present the results for 2012 and 2018. The results for other years are similar, and are available on request.

We first estimate the model including digital skills as a linear variable instead of digital

Table 7: Sensitivity checks employment model

			2012		2018	
			Coeff.	se	Coeff.	se
Specification	Linear digital skills		0.0263	0.0223	0,0421**	0.021
	Only digital skills	Digital skills: level 0	0.0957	0.0662	0.1186**	0.06
		Digital skills: level 1	0.2299***	0.0522	0.2338***	0.0466
		Digital skills: level 2	0.3108***	0.0498	0.3199***	0.0445
		Digital skills: level 3	0.3302***	0.0576	0.3340***	0.0497
	Digital skills + backgr. vars	Digital skills: level 0	0.048	0.0603	0.0854	0.0584
		Digital skills: level 1	0.1295***	0.0477	0.1587***	0.0482
		Digital skills: level 2	0.1655***	0.0464	0.1999***	0.0486
		Digital skills: level 3	0.1597***	0.0573	0.1822***	0.0539
	Digital skills + literacy and numeracy	Digital skills: level 0	0.0773	0.0639	0.1045*	0.0588
		Digital skills: level 1	0.1272**	0.0555	0.1596***	0.0508
		Digital skills: level 2	0.1452**	0.0592	0.2003***	0.0573
		Digital skills: level 3	0.1127	0.0738	0.1772***	0.0682
	Logit model	Digital skills: level 0	0.1049	0.2964	0.3219	0.285
		Digital skills: level 1	0.3643	0.279	0.6106**	0.2592
		Digital skills: level 2	0.5933*	0.3388	0.9287***	0.3254
		Digital skills: level 3	0.4952	0.5536	0.7556	0.4686
Other employment definitions	Employed for more than 6 months per year	Digital skills: level 0	0.0624	0.0569	0.0817	0.0604
		Digital skills: level 1	0.1093**	0.0508	0.1377***	0.0502
		Digital skills: level 2	0.1192**	0.057	0.1778***	0.0572
		Digital skills: level 3	0.0898	0.074	0.1506**	0.0675
	Employed in the middle of a year (july)	Digital skills: level 0	0.0372	0.0572	0.092	0.059
		Digital skills: level 1	0.0920*	0.0508	0.1429***	0.0492
		Digital skills: level 2	0.1042*	0.0565	0.1773***	0.0551
		Digital skills: level 3	0.0724	0.0741	0.1394**	0.0643
	Employment (based on income register data)	Digital skills: level 0	0.0562	0.0564	0.0906	0.0612
		Digital skills: level 1	0.1019**	0.0494	0.1174**	0.0549
		Digital skills: level 2	0.1329**	0.0627	0.1522**	0.0642
		Digital skills: level 3	0.0945	0.0783	0.1256	0.0864
Sample	All ages	Digital skills: level 0	0.0800*	0.043	0.0518	0.0332
		Digital skills: level 1	0.1426***	0.0357	0.0921***	0.0317
		Digital skills: level 2	0.1559***	0.0441	0.1145***	0.0407
		Digital skills: level 3	0.1044*	0.0544	0.0952**	0.0477
	No imputation	Digital skills: level 0	-	-	-	-
		Digital skills: level 1	0.0792	0.0519	0.0799	0.05
		Digital skills: level 2	0.1122*	0.0589	0.1261**	0.0622
		Digital skills: level 3	0.1015	0.0751	0.1128	0.0756
	Full imputation	Digital skills: level 0	0.0424	0.0611	0.0804	0.0592
		Digital skills: level 1	0.095*	0.0514	0.1250**	0.0498
		Digital skills: level 2	0.1100**	0.0548	0.1458***	0.0539
		Digital skills: level 3	0.0849	0.0692	0.1110*	0.0628
		Prefer paper test	-0.018	0.0648	-0.0304	0.0665

skill dummies. For the first years after the PIAAC survey, the digital skills have no significant coefficient, but in later years the coefficient becomes significant. This is not surprising, as the results from the main model show that the effect of digital skills is nonlinear. As the effect sizes are larger in later years, even the linear model estimates a significant coefficient.

Next, we estimate three models with 1) only digital skills as explanatory variable, 2) digital skills and the background variables as explanatory variables, and 3) digital skills and literacy and numeracy as explanatory variables. When only digital skills are included, the coefficients are two to three times larger than in the main model as they take up the additional effects of literacy, numeracy and the background variables. When either the background variables or literacy and numeracy are added to the model, the coefficients for digital skills decrease, but are still higher than in the main model.[21]

As the dependent variable is a dummy, we also estimated a logit model, instead of the linear probability model in the main specification. The estimated coefficients have the expected sign, and have the same pattern over the skill levels as the main model. However, the significance is less. This is common in logit models as the nonlinear specification makes the estimates less precise.

In the main specification, we use employment in December as dependent variable, based on the socio-economic data from CBS. Other operationalizations of employment are possible. We also estimated the model using employed for more than six months in a year as dependent variable and using employment in July as dependent variable. Both models give very similar results as the main specification.

Thus far, we based our employment definition on the socio-economic data, but persons classified as not employed in the socio-economic data might still have a wage in the wage dataset, while persons classified as employed might be missing in the wage dataset. There are at least three reasons for this discrepancy. First, the wage data only contain salaried workers and do not contain information on e.g. self-employed persons.[22] Second, the socio-economic data categorize persons based on the main source of income, while the wage data also include minor side jobs. Third, we aggregate the wage at a yearly level,

[21] We also estimated the model leaving out only literacy or only numeracy. The results are very similar to the model where we include all three skills. Results are available on request.

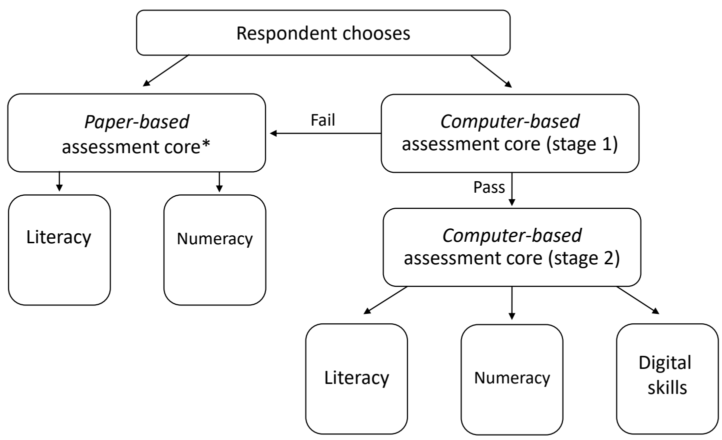
[22] In our dataset, approximately 18 percent of the active persons is self-employed at some moment in 2012-2018. This roughly matches yearly Eurostat figures on the Netherlands (16 percent in 2018, see Eurostat (2019)).

using only those months in which a wage was earned, while the employment variable is based on the employment status in December. This causes a discrepancy if someone is temporarily unemployed at the end of the year. As a sensitivity check, we construct a dummy variable that indicates whether, according to the wage dataset, someone had a wage from a salaried job at some point during the year. When we estimate a model using this dummy as dependent variable, we again find very similar results.

As a final sensitivity check, we estimate the model for different samples. When we include all ages, the main conclusions do not change. When we exclude the persons who failed the basic ICT test, and use persons with level 0 as reference category, the coefficients tend to become slightly smaller. This is not surprising, as the new reference category, level 0, has a higher employment share than the persons who failed the basic ICT test. When we add the persons who prefer a paper test to the main model, we find that those persons are very similar to the persons who failed the basic ICT test. The other results again are robust to this addition.

10.4 Descriptives

Figure 8: Structure of PIAAC survey



* We can distinguish between respondents who refused to do the Computer-based assessment core or who failed the Computer-based assessment core

Source: Derived from Figure 1.4 from OECD (2016).

Table 8: Variable descriptions

Variable	Description	Data source
Wage	Average gross hourly wage (in Euro) in year (2012-2019)[a]	CBS
Log wage growth	Difference between ln(wage 2019) and ln(wage 2012)	CBS
Wage (self-reported)	Gross hourly wage (in Euro) as reported by the PIAAC respondent	PIAAC
Employed	Categorical variable: 1 if employed in december of each year (2012-2018), 0 otherwise[b]	CBS
Employed (self-reported)	Categorical variable: 1 if the PIAAC respondent indicates (s)he is employed, 0 otherwise	PIAAC
Active	Categorical variable: 1 if employed in at least one month during 2012-2018, 0 otherwise	CBS
Contract type	Contract type in December of each year (2012-2018); categorical variable: fixed, temporary or not applicable	CBS
Unemployment spell	Categorical variable: 1 if received unemployment benefits or social welfare benefits in at least one month during 2012-2018	CBS
Digital skills	Measured by ten plausible values, on a scale between zero and 500	PIAAC
Numeracy skills	Measured by ten plausible values, on a scale between zero and 500	PIAAC
Literacy skills	Measured by ten plausible values, on a scale between zero and 500	PIAAC
Age	Age in years (in 2012)	PIAAC
Gender	Categorical variable: male or female	PIAAC
Education	Categorical variable: low (ISCED 2 and lower or less than two years of ISCED 3c), medium (ISCED 3 and 4) and high level (ISCED 5 and higher) and other (foreign qualification)	PIAAC
Children	Categorical variable: Has children (=1) and has no children (=0)	PIAAC
Born in NL	Categorical variable: Is born in the Netherlands (=1) and is not born in the Netherlands (=0)	PIAAC
Respondent type	Categorical variable: regular respondent (digital skills are measured in a regular way), respondent failed basic computer skills test or respondent prefers paper test	PIAAC

[a] Note that the hourly wage might differ per month and if a person is temporarily jobless some months will have missing observations. Moreover, a person might have more than one job at the same time. We average the hourly wages from all available months and jobs, weighted by the number of hours worked in each month and job.
[b] If a person is jobless in December, we take the contract type of the latest available month of the year. If a person has multiple jobs in December, we take the contract type of the job that has most hours.

Table 9: Frequency distribution of digital, numeracy and literacy skills (for different samples, in percentages)

Sample type	(1) Main	(2) All ages	(3) No imputation	(4) Full imputation
Age restriction (30-55 years)	Yes	No	Yes	Yes
Impute those who failed basic ICT test	Yes	Yes	No	Yes
Impute those who prefer a paper test	No	No	No	Yes
Sample size				
n	2,610	4,855	2,477	2,712
Skills				
Digital skills				
Level 0	11.809	12.921	12.564	11.308
Level 1	35.228	34.533	37.479	33.734
Level 2	38.854	37.177	41.336	37.206
Level 3	8.103	8.222	8.621	7.759
Failed basic ICT test	6.005	7.147	-	5.751
Prefer a paper test	-	-	-	4.242
Numeracy skills				
Level 0	2.981	2.985	1.399	3.542
Level 1	8.527	9.381	7.216	8.909
Level 2	25.365	28.006	25.392	25.642
Level 3	41.905	41.066	43.717	41.258
Level 4	19.391	17.062	20.373	18.876
Level 5	1.830	1.499	1.904	1.773
Literacy skills				
Level 0	2.234	2.268	1.142	2.524
Level 1	7.560	8.536	6.189	8.192
Level 2	23.958	26.283	23.575	24.142
Level 3	44.209	43.174	45.980	43.623
Level 4	20.758	18.348	21.777	20.188
Level 5	1.279	1.391	1.339	1.331

Table 10: Frequency distribution of background variables (for different samples, in percentages)

Sample type	(1) Main	(2) All ages	(3) No imputation	(4) Full imputation
Age restriction (30-55 years)	Yes	No	Yes	Yes
Impute those who failed basic ICT test	Yes	Yes	No	Yes
Impute those who prefer a paper test	No	No	No	Yes
Sample size				
n	2,610	4,855	2,477	2,712
Background variables				
Average age	43.24	41.29	43.09	43.36
Male	47.78	49.52	47.96	47.71
Female	52.22	50.48	52.04	52.29
Education				
Less than high school	22.95	29.89	20.79	23.97
High school	38.39	38.52	39.32	38.35
Above high school	37.78	30.92	39.20	36.80
Foreign qualification	0.88	0.68	0.69	0.88
Children				
No children	21.95	37.63	21.76	21.94
Yes	78.05	62.37	78.24	78.06
Born in the Netherlands				
Yes	90.11	91.45	91.48	89.34
No	9.89	8.55	8.52	10.66

10.5 Regression tables

Table 11: Regressions log(wage) in 2012-2019 and wage growth

Dependent variable log (hourly wage)	(1) **2012**	(2) **2013**	(3) **2014**	(4) **2015**	(5) **2016**	(6) **2017**	(7) **2018**	(8) **2019**	(9) **growth**
Digital skills	0.0359**	0.0534***	0.0485***	0.0415**	0.0515***	0.0363**	0.0434**	0.0378**	0.0065
	(0.0164)	(0.0160)	(0.0158)	(0.0207)	(0.0172)	(0.0185)	(0.0187)	(0.0174)	(0.0181)
Numeracy skills	0.0306	0.0336	0.0387	0.0473*	0.0369	0.0542**	0.0408	0.0386	-0.0049
	(0.0255)	(0.0235)	(0.0236)	(0.0267)	(0.0250)	(0.0274)	(0.0260)	(0.0239)	(0.0188)
Literacy skills	0.0359	0.0214	0.0219	0.0216	0.0283	0.0266	0.0338	0.0259	0.0079
	(0.0247)	(0.0233)	(0.0240)	(0.0274)	(0.0271)	(0.0299)	(0.0303)	(0.0269)	(0.0198)
log (age)	0.3948***	0.3583***	0.3323***	0.2939***	0.2356***	0.1825***	0.1776***	0.1232**	-0.2454***
	(0.0452)	(0.0475)	(0.0474)	(0.0555)	(0.0491)	(0.0487)	(0.0472)	(0.0512)	(0.0341)
Education: medium level	0.1096***	0.1016***	0.1080***	0.1233***	0.0920***	0.1011***	0.1030***	0.1089***	-0.0089
	(0.0177)	(0.0188)	(0.0175)	(0.0233)	(0.0206)	(0.0207)	(0.0211)	(0.0195)	(0.0184)
Education: high level	0.4214***	0.4116***	0.4218***	0.4445***	0.4233***	0.4299***	0.4360***	0.4432***	0.0177
	(0.0231)	(0.0223)	(0.0225)	(0.0282)	(0.0241)	(0.0240)	(0.0268)	(0.0248)	(0.0238)
Education: other	0.1533*	0.1984*	0.2083**	0.2389**	0.2021*	0.1691*	0.2294**	0.3208***	0.1768**
	(0.0881)	(0.1076)	(0.1060)	(0.1180)	(0.1062)	(0.0910)	(0.1011)	(0.0971)	(0.0895)
Not born in NL (=1)	-0.0601*	-0.0697**	-0.0661**	-0.0647**	-0.0846**	-0.0865***	-0.0732**	-0.1085***	-0.0116
	(0.0315)	(0.0274)	(0.0262)	(0.0289)	(0.0344)	(0.0326)	(0.0341)	(0.0296)	(0.0222)
Female (=1)	-0.0421	-0.0421	-0.0504	-0.0531	-0.0505	-0.0859**	-0.0779*	-0.0145	0.0278
	(0.0324)	(0.0311)	(0.0317)	(0.0350)	(0.0377)	(0.0387)	(0.0434)	(0.0355)	(0.0217)
Has children (=1)	0.1673***	0.1610***	0.1729***	0.1674***	0.1834***	0.1483***	0.1573***	0.1893***	0.0239
	(0.0238)	(0.0240)	(0.0266)	(0.0295)	(0.0258)	(0.0289)	(0.0281)	(0.0286)	(0.0187)
Female*Has children (interaction)	-0.1838***	-0.1646***	-0.1618***	-0.1617***	-0.1656***	-0.1199***	-0.1225**	-0.2075***	-0.0215
	(0.0360)	(0.0367)	(0.0367)	(0.0387)	(0.0407)	(0.0427)	(0.0476)	(0.0386)	(0.0229)
Adjusted R-squared	0.4040***	0.4014***	0.4150***	0.3600***	0.4010***	0.4021***	0.3934***	0.4074***	0.0313***
	(0.0193)	(0.0197)	(0.0195)	(0.0267)	(0.0187)	(0.0187)	(0.0195)	(0.0154)	(0.0086)
Constant	1.1692***	1.3217***	1.4274***	1.5735***	1.8296***	2.0635***	2.0929***	2.3031***	1.0371***
	(0.1647)	(0.1731)	(0.1753)	(0.2110)	(0.1809)	(0.1759)	(0.1711)	(0.1853)	(0.1294)
Observations	2,120	2,088	2,060	2,038	2,003	1,912	1,801	1,971	1,884

Notes: Standard errors in parentheses. Reference category: low education level.
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Baseline employment regressions 2012-2018

Dependent variable Employed at the end of year	(1) **2012**	(2) **2013**	(3) **2014**	(4) **2015**	(5) **2016**	(6) **2017**	(7) **2018**
Digital skills: level 0	0.0447	0.0143	0.0174	0.0565	0.0858	0.0541	0.0851
	(0.0606)	(0.0591)	(0.0589)	(0.0587)	(0.0629)	(0.0614)	(0.0591)
Digital skills: level 1	0.0966*	0.0737	0.0833	0.0742	0.1130**	0.1087**	0.1405***
	(0.0511)	(0.0494)	(0.0518)	(0.0501)	(0.0521)	(0.0510)	(0.0500)
Digital skills: level 2	0.1115**	0.0987*	0.1058*	0.1004*	0.1367**	0.1469**	0.1689***
	(0.0551)	(0.0580)	(0.0574)	(0.0524)	(0.0591)	(0.0575)	(0.0553)
Digital skills: level 3	0.0869	0.0711	0.0835	0.0814	0.1002	0.1143*	0.1400**
	(0.0701)	(0.0763)	(0.0763)	(0.0698)	(0.0692)	(0.0683)	(0.0647)
Numeracy skills	0.0341	0.0359*	0.0343*	0.0253	0.0273	0.0168	0.0128
	(0.0221)	(0.0207)	(0.0204)	(0.0184)	(0.0184)	(0.0198)	(0.0182)
Literacy skills	-0.0000	0.0038	-0.0031	0.0185	0.0063	0.0008	0.0072
	(0.0248)	(0.0257)	(0.0241)	(0.0216)	(0.0235)	(0.0269)	(0.0238)
Age: 36-40 years	0.0190	0.0239	0.0163	-0.0006	-0.0104	0.0105	0.0121
	(0.0234)	(0.0230)	(0.0255)	(0.0272)	(0.0272)	(0.0246)	(0.0218)
Age: 41-45 years	-0.0036	-0.0051	-0.0185	0.0037	-0.0214	-0.0302	-0.0380*
	(0.0205)	(0.0245)	(0.0247)	(0.0259)	(0.0251)	(0.0237)	(0.0225)
Age: 46-50 years	0.0079	0.0315	0.0230	0.0085	0.0070	-0.0124	-0.0133
	(0.0244)	(0.0244)	(0.0269)	(0.0285)	(0.0275)	(0.0242)	(0.0230)
Age: 51-55 years	-0.0216	-0.0165	-0.0455	-0.0621*	-0.0748**	-0.1057***	-0.1041***
	(0.0301)	(0.0300)	(0.0296)	(0.0326)	(0.0311)	(0.0283)	(0.0280)
Education: medium level	0.0688***	0.0644**	0.0709***	0.0566**	0.0582**	0.0532**	0.0542**
	(0.0259)	(0.0254)	(0.0268)	(0.0258)	(0.0261)	(0.0265)	(0.0258)
Education: high level	0.1168***	0.1072***	0.1100***	0.0922***	0.1005***	0.1067***	0.1025***
	(0.0259)	(0.0242)	(0.0256)	(0.0255)	(0.0242)	(0.0243)	(0.0249)
Education: other	0.1782**	0.0994	0.0922	0.1134	0.1158	0.0961	0.0890
	(0.0806)	(0.0835)	(0.0931)	(0.0964)	(0.0968)	(0.0949)	(0.1025)
Not born in NL (=1)	-0.1765***	-0.1444***	-0.1526***	-0.1430***	-0.1418***	-0.1286***	-0.0964***
	(0.0347)	(0.0332)	(0.0330)	(0.0327)	(0.0346)	(0.0336)	(0.0341)
Female(=1)	-0.0257	-0.0453	-0.0666**	-0.0471	-0.0685*	-0.0767**	-0.1106***
	(0.0323)	(0.0319)	(0.0308)	(0.0341)	(0.0362)	(0.0351)	(0.0328)
Has children (=1)	0.0700***	0.0713***	0.0724***	0.0992***	0.0831***	0.0809***	0.0402*
	(0.0213)	(0.0229)	(0.0236)	(0.0274)	(0.0259)	(0.0245)	(0.0218)
Female*Has children (interaction)	-0.0971***	-0.0740**	-0.0685**	-0.0871**	-0.0729*	-0.0656*	-0.0154
	(0.0357)	(0.0345)	(0.0337)	(0.0384)	(0.0383)	(0.0376)	(0.0371)
Adjusted R-squared	0.1209***	0.1125***	0.1143***	0.1088***	0.1084***	0.1098***	0.1034***
	(0.0168)	(0.0153)	(0.0152)	(0.0143)	(0.0138)	(0.0145)	(0.0138)
Constant	0.6813***	0.6860***	0.6883***	0.6807***	0.6686***	0.6847***	0.6915***
	(0.0542)	(0.0567)	(0.0545)	(0.0581)	(0.0591)	(0.0565)	(0.0533)
Observations	2,608	2,602	2,595	2,588	2,582	2,576	2,563

Notes: Standard errors in parentheses. Reference categories: digital skills- failed basic ICT test, low education level and individuals aged 30-35 years.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13: Regressions labor force participation, unemployment spells and contract type

Dependent variable	(1) **Active in 2012-2018**	(2) **Employed in December 2012**	(3) **Had unemploy-ment spell**	(4) **Fixed contract December 2012**
Conditional on active in 2012-2018	No	Yes	Yes	No
Digital skills: level 0	0.0861*	-0.0189	0.0678	-0.0533
	(0.0453)	(0.0617)	(0.0585)	(0.0747)
Digital skills: level 1	0.1220***	0.0027	0.0750	0.0113
	(0.0431)	(0.0488)	(0.0558)	(0.0633)
Digital skills: level 2	0.1219***	0.0142	0.0718	0.0414
	(0.0471)	(0.0495)	(0.0595)	(0.0686)
Digital skills: level 3	0.0900	0.0146	0.1062	0.0143
	(0.0558)	(0.0581)	(0.0733)	(0.0904)
Numeracy skills	0.0159	0.0263	-0.0442*	0.0051
	(0.0158)	(0.0200)	(0.0267)	(0.0246)
Literacy skills	0.0165	-0.0145	0.0327	0.0007
	(0.0184)	(0.0200)	(0.0274)	(0.0269)
Age: 36-40 years	0.0063	0.0153	-	-
	(0.0166)	(0.0213)	-	-
Age: 41-45 years	-0.0220	0.0160	-	-
	(0.0165)	(0.0188)	-	-
Age: 46-50 years	-0.0041	0.0131	-	-
	(0.0195)	(0.0215)	-	-
Age: 51-55 years	-0.0478**	0.0278	-	-
	(0.0234)	(0.0235)	-	-
Log (age)	-	-	-0.1126**	0.2813***
	-	-	(0.0553)	(0.0611)
Education: medium level	0.0464**	0.0335*	-0.0321	0.0024
	(0.0207)	(0.0197)	(0.0250)	(0.0272)
Education: high level	0.0679***	0.0616**	-0.0867***	0.0199
	(0.0216)	(0.0240)	(0.0253)	(0.0251)
Education: other	0.0575	0.1630***	-0.2256***	-0.0443
	(0.0762)	(0.0630)	(0.0762)	(0.1049)
Not born in NL (=1)	-0.0844***	-0.1249***	0.1227***	-0.1265***
	(0.0288)	(0.0299)	(0.0351)	(0.0395)
Female(=1)	-0.0680***	0.0387	-0.0004	0.0449
	(0.0257)	(0.0249)	(0.0406)	(0.0473)
Has children (=1)	0.0282*	0.0445**	-0.0924***	0.1268***
	(0.0158)	(0.0194)	(0.0289)	(0.0408)
Female*Has children (interaction)	-0.0319	-0.0753***	0.0156	-0.0942*
	(0.0295)	(0.0279)	(0.0425)	(0.0533)
Adjusted R-squared	0.1094***	0.0425***	0.0293***	0.0442***
	(0.0160)	(0.0142)	(0.0097)	(0.0135)
Constant	0.7986***	0.8425***	0.6499***	0.6521***
	(0.0473)	(0.0540)	(0.2277)	(0.2460)
Observations	2,610	2,389	2,391	2,055

Notes: Standard errors in parentheses. Reference categories: digital skills- failed basic ICT test, low education level and individuals aged 30-35 years.
*** p<0.01, ** p<0.05, * p<0.1

10.6 Additional analyses

Table 14: Comparing baseline results with results on self-reported wage as measured in PIAAC

Dependent variable log (hourly wage)	(1) **PIAAC data** **2012**	(2) **trimmed PIAAC data** **2012**	(3) **Register data (baseline)** **2012**
Digital skills	0.0554	0.0265	0.0359**
	(0.0397)	(0.0242)	(0.0164)
Numeracy skills	-0.0064	-0.0233	0.0306
	(0.0431)	(0.0323)	(0.0255)
Literacy skills	0.1009**	0.0960***	0.0359
	(0.0506)	(0.0312)	(0.0247)
log(age)	0.4798***	0.2594***	0.3948***
	(0.0868)	(0.0563)	(0.0452)
Education: medium level	0.1375***	0.0948***	0.1096***
	(0.0445)	(0.0272)	(0.0177)
Education: high level	0.3995***	0.3130***	0.4214***
	(0.0547)	(0.0310)	(0.0231)
Education: other	0.0859	0.1664	0.1533*
	(0.1444)	(0.1017)	(0.0881)
Not born in NL (=1)	0.0606	-0.0026	-0.0601*
	(0.0641)	(0.0433)	(0.0315)
Female (=1)	-0.2070***	-0.1444***	-0.0421
	(0.0702)	(0.0434)	(0.0324)
Has children (=1)	0.2284***	0.1460***	0.1673***
	(0.0405)	(0.0339)	(0.0238)
Female*Has children (interaction)	-0.5211***	-0.4185***	-0.1838***
	(0.0785)	(0.0485)	(0.0360)
Adjusted R-squared	0.3018***	0.3781***	0.4040***
	(0.0428)	(0.0225)	(0.0193)
Constant	0.7027**	1.5333***	1.1692***
	(0.3170)	(0.2074)	(0.1647)
Observations	1,780	1,521	2,120

Notes: Standard errors in parentheses. Reference category: low education level. The second model (column 2) excludes the bottom and top first percentile of the income distribution as reported in the PIAAC questionnaire. The third column returns the baseline results using register income data.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 15: Comparing baseline results with results on self-reported employment as measured in PIAAC

Dependent variable Employed (=1)	(1) **PIAAC data** **2012**	(2) **Register data (baseline)** **2012**
Digital skills: level 0	0.0563	0.0447
	(0.0559)	(0.0606)
Digital skills: level 1	0.1213**	0.0966*
	(0.0508)	(0.0511)
Digital skills: level 2	0.1424**	0.1115**
	(0.0559)	(0.0551)
Digital skills: level 3	0.1172*	0.0869
	(0.0693)	(0.0701)
Numeracy skills	0.0317	0.0341
	(0.0251)	(0.0221)
Literacy skills	-0.0035	-0.0000
	(0.0259)	(0.0248)
Age: 36-40 years	0.0266	0.0190
	(0.0269)	(0.0234)
Age: 41-45 years	0.0285	-0.0036
	(0.0235)	(0.0205)
Age: 46-50 years	0.0188	0.0079
	(0.0257)	(0.0244)
Age: 51-55 years	0.0043	-0.0216
	(0.0306)	(0.0301)
Education: medium level	0.0455*	0.0688***
	(0.0257)	(0.0259)
Education: high level	0.0754***	0.1168***
	(0.0236)	(0.0259)
Education: other	0.1622**	0.1782**
	(0.0817)	(0.0806)
Not born in NL (=1)	-0.1354***	-0.1765***
	(0.0321)	(0.0347)
Female (=1)	-0.0264	-0.0257
	(0.0330)	(0.0323)
Has children (=1)	0.0613***	0.0700***
	(0.0236)	(0.0213)
Female*Has children (interaction)	-0.1077***	-0.0971***
	(0.0358)	(0.0357)
Adjusted R-squared	0.0975***	0.1209***
	(0.0145)	(0.0168)
Constant	0.6727***	0.6813***
	(0.0567)	(0.0542)
Observations	2,608	2,608

Notes: Standard errors in parentheses. Reference category: low education level. Employment as measured in PIAAC (column 1) is set to 1 if respondents indicated to have had a paid job at the time the survey was taken.
*** p<0.01, ** p<0.05, * p<0.1