Skill-Biased Innovation Activities: Evidence from Hungarian Firms*

Attila Lindner
University College London

Balázs Muraközy
CERS, Hungarian Academy of Sciences

Balázs Reizer
CERS, Hungarian Academy of Sciences

December 2018

The most recent version of this paper can be downloaded from here.

Abstract

This paper investigates the consequences of innovation activities. We exploit a unique firm-level survey linked to employee data from Hungary that allows us to examine broadly defined innovation activities including the introduction of new products, process innovation and organizational innovation. We show that these innovation activities are skill-biased insofar as they lead to an increase both in the share of college educated workers and in their wage premium. The skill bias is not solely driven by high-novelty, R&D-based innovation, but also, to a comparable extent, by the low-novelty kind. Among low-novelty innovation types, product and process innovation are the most skill-biased, while organizational innovation is less so. These results highlight that low-novelty innovations contribute substantially to wage inequality.

keywords: skill-biased technological change, innovation, skill premia

JEL-codes: J31, J24, O30, O33

*We thank Dávid László, Hajnalka Katona and Dániel Pass for the excellent research assistance. We thank Márton Csillag, Miklós Koren, John van Reenen, Richard Blundell and seminar participants at Central European University, London School of Economics, University College London, Queen Mary University, IAB and Norwegian School of Economics for their insightful comments. We also thank the Hungarian Academy of Sciences for funding this research as part of its Momentum Grant "Firms, Strategy and Performance" and the National Research, Development and Innovation Office for funding via its K/125101 grant and the MTA Premium Postdoctoral Research Program. We are also grateful to the Central Statistical Office of Hungary for making the data available.

1 Introduction

Innovation is the main driver of economic growth. However, the process of innovation also affects the allocation of resources. The effect of innovation and technological change on income inequality is in a central focus of current policy debates. While it is widely accepted that technological change is a key driver of increasing income inequality (Acemoglu 2002, Goldin & Katz 2010), our understanding of the key mechanisms behind this relationship is quite limited, and nearly exclusively relies on proxies of innovation which measure the generation of highly novel knowledge, i.e. involving R&D or generating patents. Moreover, even regarding the skill bias of R&D, the evidence is inconclusive (Aghion et al. 2017, Bøler 2015).[1]

In fact, most innovation activity has a relatively low-novelty content. Only between 25-35 percent of process innovator firms introduced a process that was 'new to the market', and 5-25 percent of product innovator enterprises introduced products which were 'new to the world' in European countries (Figure A1).[2] On the innovation input side, the majority (50-80%) of innovators typically introduced products or processes without relying on dedicated R&D expenditures.[3] Accordingly, low-novelty innovation has the potential to play a major role in technology diffusion and thus in aggregate technological change (Mokyr 2003, Bloom et al. 2016).

Cross-country patterns suggest that innovation is related to skill demand. Figure 1, Panel A shows that there is a strong relationship between the fraction of innovating firms and the college premium among Western European countries. The positive correlation holds even if we control for the supply of college graduates and level of GDP per capita, or if we additionally include new EU members states (Table A1). At the same time, we find no clear relationship between the share of R&D conducting firms and the wage premium (Figure 1, Panel B). This evidence is indicative of a substantial role of low novelty innovation in aggregate skill demand and, in turn, inequality.

This paper investigates the relationship between firm-level skill demand, proxied by the share and wage premium of college-educated workers, and different innovation activities, which involve the introduction of production processes, products and management methods which are new for the firm but not necessarily for the market. Our focus on firms is motivated by recent evidence that highlights the crucial role firms play in explaining increasing inequality (Song et al. 2015, Card et al. 2018). Moreover, focusing on firms allows us to identify the effect of various innovation activities using difference-in-differences style estimation strategies.

A key novelty of the paper is the use of exceptionally rich microdata from Hungary on innovation, which is linked to worker-level wage and employment information. We rely on the rich information

[1] Aghion et al. (2017) finds that more R&D intensive firms pay a lower college premium, while Bøler (2015) finds that higher R&D intensity is associated with an increase in the skill ratio.

[2] For example, in France, which is at the higher end of the range, 31 percent of process innovations are new to the market and 24 percent of product innovations are new to the world. Clearly, the bulk of innovation that takes place at the firm level has relatively low-novelty value.

[3] While R&D-based innovation tends to be somewhat more frequent in countries closer to the technological frontier, it is quite prevalent even in innovation leaders (Germany: 38 percent, France: 47 percent, Finland: 64 percent). Neither does R&D expenditure dominate total innovative spending: its cost share is around 50 percent in a typical European country. Other measures of novelty suggest a similar picture.

available from the Community Innovation Survey (CIS), which employs a very inclusive definition of innovation on the one hand, and asks specific questions about the novelty value and the type of the innovation on the other. A conceptual advantage of CIS compared to other types of innovation proxies, such as R&D, is that it properly measures innovation outputs rather than inputs (Mairesse & Mohnen 2010).

Our primary focus is understanding the relationship between firm-level innovation activities and skill demand. To relate to the recent literature on firm-level wage premia, we introduce a theoretical framework where firms have wage setting power along the lines of Card et al. (2018). This framework allows us to separately investigate the relationship between innovation and both the skill ratio and the skill premium. We derive that even a Hicks-neutral shock can affect the skill ratio when firms face a different wage-setting power in the skilled and unskilled labor markets. The intuition is that following a Hicks-neutral shock, firms will hire more new employees in the more competitive market where there increased demand drives up prices to smaller extent. This implies that focusing only on the skill ratio (see e.g. Caroli & Van Reenen 2001, Bøler 2015) is not sufficient to identify whether technological change is skill biased. However, the joint investigation of the skill premium and the skill ratio allows one to identify skill-biased innovation activities.

Motivated by this theoretical framework, we investigate responses at both the quantity and wage margins. We estimate the relationship between skill ratio and innovation activities by following the identification strategy of the seminal paper of Caroli & Van Reenen (2001). This approach investigates how innovation decisions are related to subsequent long (6-year) changes in skill ratio and firm performance. This framework is not only suitable to handle unobserved firm heterogeneity but is also able to capture the long-term effects of innovation. We find that innovation, either low- or high-novelty, is associated with subsequent growth both in the employment and wage share of college-educated workers and subsequent productivity growth.

We estimate how the skill premium is related to innovation at the worker level. We implement a diff-in-diff style identification strategy where we compare workers at firms which start to innovate to workers in firms which remain non-innovative. To alleviate the concern that innovative firms pay higher wages and premia even before innovating, we match firms based on observable characteristics in the initial period. Alternatively, we focus on firms which start to innovate at different points of time, thus we only exploit variation in the timing of innovation. With both empirical strategies we find that starting to innovate is associated with a 5-8 percentage point increase in the wage premium. This evidence suggests that innovation activities are skill biased in Hungary.

The results are not driven by pre-existing trends, robust to various matching procedures, not sensitive to alternative timing assumptions and do not change when controlling for industry and occupation specific shocks. Furthermore, these results are not driven by the change in the composition of the workforce.[4]

[4] In particular, we show that the results are robust to restricting the sample to incumbent workers (who had worked in the firm for at least 24 months). Results are also unchanged when we proxy for unobserved heterogeneity by following Koren & Csillag (2017). Moreover, using a subset of individuals who we can link across years, we show that the effect of innovation on the wage premium is similar with and without individual fixed effects.

We also show that the increase in the wage premium is persistent. In other words, it does not seem to result from temporary higher efforts by college educated workers during the introduction of the innovation, but from long-lasting technological change. We also find that innovation is associated with an increased premium of nonroutine workers, but this is unrelated to the increasing college premium.

After establishing that innovation is skill biased, we study the heterogeneity of innovation along two dimensions. First, we are interested in the extent to which the novelty of the innovation is associated with skill ratio and premia. We quantify novelty in three ways: whether the innovation involved R&D, whether it was new to the market, and whether the firm itself has developed it. We find that both low- and high-novelty innovations are associated with an increase in the college premium, and that these magnitude of these changes is quite similar to one another. Given its prevalence in economy, low novelty innovation plays a larger role in explaining the skill premium in Hungary than high novelty innovation.

Second, we distinguish between technological and organizational innovation, and between process and product innovation. While all these types of innovation appear to increase the skill ratio to similar extent, we find that the skill premium is mainly driven by technological innovation, especially process innovation. These evidence suggest that the process of technological adoption is skill biased in Hungary while the evidence regarding organizational innovation is less conclusive.

Our paper is related to several strands of literature.

First, we contribute to the literature that explains the evolution of wage inequality with skill biased technological change (Acemoglu 2002, Goldin & Katz 2010). Instead of focusing on specific technologies, such as computers (Autor et al. 1998) and broadband internet (Akerman et al. 2015), or high novelty innovation, such as R&D (Aghion et al. 2017, Bøler 2015) and patents (Kline et al. 2018), here we consider all innovation activities and technology adoption. This more inclusive investigation can capture a much larger share of the technological change taking place in the economy.

Our paper also contributes to the literature that directly considers firm-level skill demand and technological change or innovation (Caroli & Van Reenen 2001, Bresnahan et al. 2002, Abowd et al. 2007). These studies usually rely on relatively small cross-sectional surveys measuring specific innovation activities or the implementation of specific technologies. In contrast, our data includes five repeated waves of a large-scale innovation survey, each of which covering a large set of firms (around 5000 firms), and provides consistent measures for various types of innovation activities over time (and across countries). The panel dimension of our survey allows us to implement empirical strategies (e.g. using matching or switching sample) that provide more credible estimates on the effect of innovation on skill demand. Furthermore, we emphasize in our theoretical approach that an increase in the skill ratio - the focus of most of the existing literature - may not be an unequivocal sign of skill biased technological change when labor markets are not competitive. As a result, we investigate both the quantity and wage margin.

The innovation activities we study are strongly linked to the question of how firms tap into global knowledge sources and react to global competition. A strand of literature investigates international technology and knowledge spillovers and their consequences (Coe & Helpman 1995, Bayoumi et al. 1999, Keller 2004). Another strand of literature, building on the seminal paper of Acemoglu (2003), investigates whether and how firms upgrade their technology as a reaction to opportunities and shocks created by trade. One channel is that trade liberalization provides export opportunities, and, therefore a potentially higher returns to technology upgrading (Costantini & Melitz 2008). Evidence for such trade-induced technological change is provided by, for example, Bustos (2011b), while Verhoogen (2008) and Bustos (2011a) also show that opening to trade is associated with SBTC. Another mechanism between globalization and SBTC is the reduced cost of importing technology embedded in machines (Caselli 2014, Koren & Csillag 2017) or inputs (Kasahara et al. 2016, Akhmetova & Ferguson 2015). While this paper does not directly rely on trade shocks, the innovation activities we observe may largely be driven by international knowledge flows, either embedded in capital goods or driven by global competition. Importantly, in contrast to some of these studies, we directly observe the technology adoption decision at the firm level and do not have to rely on proxies, such as innovation inputs.

In what follows, Section 2 describes briefly the conceptual framework we rely on. Section 3 describes our data sources and Section 4 discusses our empirical strategy. Section 5 presents our results, and Section 6 concludes.

2 Conceptual framework

This section introduces a conceptual framework that underline our empirical approach. The main aim is to present a framework which provides predictions on how Hicks neutral and skill biased technological change affects the skill ratio and the skill premium. To this end, we introduce some wage setting power at the firm-level as in Card et al. (2018).

Following Card et al. (2018) we assume that there are J firms and two types of workers: lower-skilled (L) and higher-skilled (H). Each firm $j \in 1, ..., J$ posts a pair (w_{Lj} , w_{Hj}) of skill-specific wages that all workers costlessly observe. For worker i in skill group $S \in L, H$, the indirect utility of working at firm j is

$$\ln u_{iSj} = \beta_S \ln w_{Sj} + a_{Sj} + \epsilon_{iSj}$$

where a_{Sj} is a firm-specific amenity common to all workers in group S and ϵ_{iSj} captures idiosyncratic preferences for working at firm j. We assume that the ϵ_{iSj} are independent draws from a type I Extreme Value distribution. Card et al. (2018) derive that if J is large the approximate firm-specific supply functions are:

$$\ln L_j(w_{Lj}) = \ln(L\lambda_L) + \beta_L \ln w_{Lj} + a_{Lj}$$

$$\ln H_j(w_{Hj}) = \ln(H\lambda_H) + \beta_H \ln w_{Hj} + a_{Hj}$$

where (λ_H, λ_L) are constants common to all firms in the market. Note that as $\beta_L, \beta_H \to \infty$, these supply functions become perfectly elastic and we approach a competitive labor market.

Firms have production functions of the form:

$$Y_j = TFPR_j \left[(1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}} + \theta_j H_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

where σ is the elasticity of subsitution between high- and low-skilled labor. $TFPR_j$ is the Hicks-neutral (revenue) productivity. θ_j dictates the relative importance of high-skilled labor in production. Skill biased technological change is therefore captured by an increase in θ_j.

Firms minimize their cost given the labor supply equation they face. It is easy to show that the FOCs are equivalent to:

$$\ln \frac{w_{Hj}}{w_{Lj}} = \ln \frac{\beta_H}{1+\beta_H} - \ln \frac{\beta_L}{1+\beta_L} + \ln \frac{\theta_j}{1-\theta_j} - \frac{1}{\sigma} \ln \frac{H_j}{L_j} \tag{1}$$

The equation highlights that the firm-level wage premium depends on the β_H and β_L (the extent of which the labor markets are competitive), the skilled biased technological part, θ_j, and the high-skilled/low-skilled ratio. It is also clear that, conditional on this latter ratio, only skill biased innovation affects directly the wage premium, while Hicks-neutral innovation is only related to the premium indirectly via the skill ratio.

In general, there is no closed-form expression for for the high/low-skill ratio. However, it is worth considering a special case where $\sigma = \infty$ and so there is a perfect substitution between high-skilled and low-skilled labor. In that case the wage premium and high/low-skilled ratio is given by the following equations:

$$\ln \frac{w_{Hj}}{w_{Lj}} = \ln \frac{\beta_H}{1+\beta_H} - \ln \frac{\beta_L}{1+\beta_L} + \ln \frac{\theta_j}{1-\theta_j} \tag{2}$$

$$\ln \frac{H_j}{L_j} = C + \frac{a_{Hj}}{a_{Lj}} + \beta_H \ln \theta_j - \beta_L \ln \frac{1}{1-\theta_j} + (\beta_H - \beta_L) \ln TFPR_j \tag{3}$$

Three things should be noted. First, unlike the general case, the wage premium ($\ln \frac{w_{Hj}}{w_{Lj}}$) does not depend on the skill ratio. This highlights that whatever the effect of innovation on the skill ratio ($\ln \frac{H_j}{L_j}$) is, the wage premium is unaffected by that. Consequently, examining the effect on wage premium is sufficient to determine whether innovation is skill-biased or not.

Second, naturally, the skill ratio is directly related to the the skill biased technological parameter, θ_j. More surprisingly, if firms' wage setting power differ in the high-skilled and low-skilled labor markets (formally if $\beta_H \neq \beta_L$), the skill ratio is also affected by Hicks-neutral shocks (changes in $TFPR_j$). For instance, if the high-skilled labor market is more competitive than the low-skilled labor market ($\beta_H > \beta_L$), firms hire more high-skilled workers following a positive shock in $TFPR_j$ because their increased demand drives up wages only to a smaller extent on that labor market.

Third, since the skill ratio can also be affected by Hicks-neutral shocks, a positive relationship between skill ratio and innovation does not necessarily reflect skill biased technological change. In contrast, the positive relationship between innovation and the skill premium provide a clear evidence for skill-biased technical change. This highlights a major limitation of the existing studies that often focus on changes in the skill ratio and ignore changes in the skill premium (see e.g. Caroli & Van Reenen 2001, Bresnahan et al. 2002, Abowd et al. 2007).

3 Data

Our work is based on three data sources: the Community Innovation Survey, the Structure of Earnings Survey and Balance Sheet data.

3.1 Innovation data

The first source is the Hungarian version of the Community Innovation Survey (CIS), conducted in a harmonized way in European Union member states. The richness of Community Innovation Survey (CIS) survey is exploited in the recent literature to estimate the effect of various types of innovation on firm performance (Crépon et al. 1998, Griffith et al. 2006), but so far no paper aimed to assess the relationship between skill demand and innovation.

The survey is bi-annual and covers a representative sample of manufacturing and service firms in the economy. Its questions always refer to the previous 3 years. In this paper we use six waves of the CIS survey from the period between 2004 and 2014. The sample size has been progressively increasing from about 4,000 firms in 2004 to more than 7,000 in 2014 (Table A2).

Most importantly for our purposes, the CIS asks detailed questions on the innovative activities of the firm including process, product and organizational innovations. Innovation here is defined very broadly. Namely, it is defined as the introduction of products/technologies which are new or significantly modified from the viewpoint of the firm, but are not necessarily new for the market. This enables one to capture many types of innovations, ranging from adoption of technologies to creating radically new knowledge via research or introducing products which are new to the world.

Importantly, the innovation definitions in the CIS are strongly grounded in innovation theory. Innovation, as defined by Schumpeter, means ”novel combinations of knowledge, resources etc. subject to attempts at commercialization” (Fagerberg 2007). According to this definition, R&D in itself is not innovation, but one of the inputs of innovation. Patents, while outputs of the innovation process, are very restrictive compared to the more general Schumpeterian definition. These distinctions also hint at substantial timing differences between research, invention, patents, and the actual commercialization of research results. Innovation surveys have been developed with these distinctions in mind, defining innovation according to the Schumpeterian framework cited above.

The database allows us to investigate the heterogeneity of innovation in a number of dimensions. First, it distinguishes between different types of innovation. We will rely on three main categories: product, process, and organization. Based on the CIS's categorization, will call the first two types *technological innovation*, while organization innovation is a type of non-technological innovation. *Product* *innovation* includes both product and services innovation, and is defined as 'the market introduction of a new or significantly improved good or service with respect to its capabilities, user friendliness, components or sub-systems.' A *process innovation* is defined as 'the implementation of a new or significantly improved production process, distribution method, or supporting activity.' An *organizational innovation* 'is a new organizational method in your enterprise's business practices (including knowledge management), workplace organization or external relations that has not been previously used by your enterprise.' These carefully drafted definitions have been developed by extensive work after a number of pilot surveys by Eurostat, to make sure that the results are comparable across countries and time periods.[5]

The CIS asks detailed questions about the novelty of the innovation. We measure the *novelty* value of the innovation by three dummies. One of the questions refers to whether the firm conducted in-house R&D, defined as 'research and development activities undertaken by your enterprise to create new knowledge or to solve scientific or technical problems'. We code this R&D dummy to one if the firm reports a positive in-house R&D spending. Second, for product and process innovations the Survey asks whether it was new for the market (for process innovation) or new to the country (product innovation). We define a dummy which takes the value of 1 if the firm answered 'yes' to any of these questions. Third, the survey asks about whether the innovation was developed by the firm (either single-handedly or together with other firms or institutions) or it was adopted (either with or without modifications).[6] We create a dummy which shows whether the firm reports a product and process innovation which was developed by the firm. Note that these novelty variables are only defined for (technologically) innovative firms.

Importantly, the survey design can be best described as a repeated cross-sectional one. As a result, it does not aim at surveying the same firms wave-by-wave. Nonetheless, a number of firms are observed multiple times and we can use the unique firm identifiers to follow them across the waves. A key issue with modeling the effects of innovation is timing. First, each wave of the CIS refers to innovation activities in the previous 3 years, therefore one cannot be sure when exactly the innovation took place. Second, innovation is a long-term investment, and its effects are unlikely to show up immediately. Because of these two reasons, it is unlikely that one can estimate the effect of innovation by a sharp event-study design.

We define the main innovation variable in the following way. A firm is considered to be innovative in the CIS wave conducted in year τ if, according to that CIS wave, it has undertaken product, process or organizational innovation. In the individual analysis, we define a firm as innovative in year t if it was innovative either in the corresponding wave ($\tau = t$ or $\tau = t + 1$) or one of the two previous waves. The motivation to do so is that innovation is likely to have effects beyond the horizon of each CIS wave. Importantly, this assumption does not affect the main results, as we show in Table A7.

[5]The definitions come from the CIS 2012 Questionnaire, available at: https://ec.europa.eu/eurostat/web/microdata/community-innovation-survey

[6]We rely on the question: 'Who developed these product/process innovations?' with the possible answers: 'i) Your enterprise by itself; ii) Your enterprise together with other enterprises or institutions; iii) Your enterprise by adapting or modifying processes originally developed by other enterprises or institutions; iv) Other enterprises or institutions'. We categorize the first two as developed by the firm. Note that it is possible that a firm develops an innovation without formal R&D and *vice versa*.

3.2 Structure of Earnings Survey

The Structure of Earnings Survey (Bértarifa) database is a yearly worker-level survey, which includes information on the demographic variables, including schooling, job characteristics and on the wage of workers earned in May. This database samples firms with less than 50 employees but collects information on all employees of these firms. For larger firms, it collects data on a representative sample of employees.

These data are available for each year between 2000 and 2014. The number of observations for employees of business-sector firms is between 120 and 170 thousand per year. Importantly, the dataset is repeated cross-sectionally at the worker level and it is not possible to perfectly link employees across waves.

The key variable from this survey is the *college* dummy, representing whether the worker is a college or university graduate (ISCED 2011 levels 5 to 8). Our main specifications will estimate the difference in the college wage premium between innovative and non-innovative firms. The results are similar when high-skilled workers are defined based on their occupation.[7]

This survey also includes detailed data on earnings. We measure wages by the total monthly compensation including regular and bonus payments to reflect all sources of income differences between workers. We also use information on the base wage of the worker, and whether the worker received any flexible wage elements. We also observe whether the worker is an incumbent, i.e. hired in the last year, and we also observe the tenure.

3.3 Balance sheet data

For important firm-level variables, including the number of employees, industry classification, ownership and key financial variables, we rely on administrative data collected for tax purposes by the Tax Office. This database includes balance sheet and profit and loss statements from all double-entry bookkeeping enterprises in Hungary. After deflating the nominal variables, we estimate TFP from these data by relying on the methods of Ackerberg et al. (2015) and Levinsohn & Petrin (2003). Our results are mostly robust to using alternative measures of productivity.

The dataset is available between 2000 and 2014 and it includes about 400-500 thousand firms per year.

[7] We also use additional skill level proxies as controls. Our base category is workers with primary education. Second, we define *high-school* as a dummy representing whether the employee has finished high-school requiring a school-leaving maturity exam (érettségi). Third, we define *vocational* skills for employees with vocational training, but no maturity exam.

3.4 Linking the databases

The three databases can be linked based on unique firm identifiers.[8]. Importantly, the CIS is representative of firms with at least 10 employees and the Structure of Earnings Survey is representative of workers. As a result, the intersection of the two databases is representative of workers working for firms with at least 10 employees.[9]

Table A2 shows the number of observations in our sample. The second column shows the number of firms in the CIS, which was around 4,000 in the beginning of the period and increased to more than 7,200 by 2014. The next column shows that about 80% of these can be merged to the balance sheet data. From this, a bit less than half, altogether nearly 24,000 firm-year-level observations, representing 6,700 firms, were sampled in the Structure of Earnings Survey. The number of employees varies between 40 thousand and 64 thousand across years, with a total of more than 700 thousand individual observations in our sample.

Table A3 shows the number of observations and the number of firms conducting different types of innovation by one-digit industry in the regression sample. More than half of the observations comes from manufacturing, while trade, transportation and utilities are also well-represented.

Each type of innovation is conducted by about 20 percent of firms, with more than half of the firms not innovating. Note that one firm can conduct more than one type of innovation. Altogether 43% of firms in our sample conducted at least one type of innovation. The prevalence of innovation is the highest in the ICT sector (55%) and lowest in Construction (28%), with Manufacturing close to the average (39%).

4 Empirical approach

First, we estimate country-industry-level regressions for European countries to set the stage. Second, we use firm-level long difference regressions to estimate the relationship between innovation and firm-level outcomes. Because of worker heterogeneity, such regressions are less suitable to investigate the wage effects of innovation, therefore, we rely on worker-level regressions to estimate the skill premium of innovative firms. We test for the heterogeneous effects of different types of innovation in both frameworks. Finally, we use a simple decomposition to quantify the importance of low- and high-novelty innovation in the aggregate cross-sectional skill premium.

[8]Given the confidential nature of these data, the merged database can be used solely on-site in the Central Statistical Office of Hungary

[9]For example, the share of workers working for firms in the different size categories in the matched dataset is very similar to the share of workers working in each size category according to the balance sheet data.

4.1 Country-industry-level regressions

We start our investigations at the country-industry level to study whether broadly defined innovation is correlated with skill demand. For this exercise, we use data from the 2010 CIS at the 1-digit industry-country level, and link it to information on the share and premium of college educated workers from the 2010 and 2014 waves of the Structure of Earning Survey.[10] These data are from Eurostat's webpage.[11]

Our question is whether the college share or the college premium increased faster in industries/countries with a higher share of innovative firms. Our empirical strategy follows Machin & Van Reenen (1998) with regressing 4-year change in skill demand on proxies of technological change, the share of innovative firms in our case.[12] In particular, we run regressions of the type:

$$\Delta y_{cst} = \beta * innovation_{cst} + \delta * y_{cst} + \eta_{ct} + \zeta_{st} + \epsilon_{cst} \tag{4}$$

where c indexes countries, s sectors (1-digit) and t time periods. Δy_{cst} is the long difference, the change of y_{cst} between years t and $t+4$. η_{ct} are country fixed effects, while ζ_{st} are sector fixed effects.

Note that these long-difference regressions remove country-industry fixed effects and identify only from changes in skill demand. Country fixed effects also remove country-level shocks to skill supply or general economic conditions. In some specifications we also include industry fixed effects to filter out industry-level shocks.

We use two dependent variables. The first one is the share of college educated workers, and the second is the college premium, or the log difference between the average wage of workers with college and non-college education, calculated from the Structure of Earnings Survey. $innovation_{cst}$ is the share of innovative firms[13], the share of firms conducting R&D or the R&D intensity of the industry.[14]

Naturally, the number of firms and employees behind the different observations varies widely. Therefore, we weight the regressions with the number of firms in the CIS in the given country-industry cell to give more weight to observations which represent an average calculated from more observations. Consequently, our approach, in line with our general focus on firms, is closest to a cross-country firm-level regression. We cluster standard errors at the country level because skill premia are likely to be strongly correlated within each country.

[10] The Structure of Earnings Survey is conducted in every 4 years in all EU countries (but every year in Hungary). Therefore, it is available for 2002, 2006, 2010 and 2014. The CIS is not available in 2002, and the 2006-2010 period may be reflect developments related to the Great Recession. Therefore, we stick to the 2010-2014 period.

[11] This matched sample includes EU28 countries (with the exception of Greece, Malta and the United Kingdom) and Norway, altogether 25 countries.

[12] Machin & Van Reenen (1998) runs a panel regression with 4-year periods, while we run the regression only on one 4-year period.

[13] Defined as conducting either product, process, organizational and marketing innovation.

[14] R&D intensity is calculated as in-house R&D expenditures over turnover for firms in the CIS sample.

4.2 Firm-level regressions

The primary aim of firm-level regressions is to estimate how innovation in a period is related to subsequent change in skill demand and productivity. We follow Caroli & Van Reenen (2001) and estimate long-difference regressions of the form:

$$\Delta y_{jt} = \beta * innovation_{jt} + \gamma * \Delta X_{jt} + \delta * y_{jt-1} + \eta_{st} + \epsilon_{jt} \tag{5}$$

where j indexes firms, t years, and y_{jt} is the variable of interest (share of high-skilled workers or productivity) and Δy_{jt} is its change between year t and $t+6$.[15] $innovation_{jt}$ is a dummy, showing whether the firm was innovative in the corresponding or the previous CIS wave. ΔX_{jt} is the long difference in value added and capital, which we include only for the college share equations.[16] y_{jt-1} is the lagged level of the y_{jt}, which controls for potential regression to the mean issues. η_{st} are industry-year fixed effects. Standard errors are clustered at the firm level.[17]

Our firm-level identification strategy is difference-in-differences. In particular, it compares (6-year changes of) outcomes of firms which did innovate in the CIS wave at the beginning of the period with firms in the same industry and initial characteristics which did not. As Caroli & Van Reenen (2001) argues, such a long difference specification is suitable to estimate the long-run effects of innovation because it differences out firm fixed effects while capturing long-term changes rather than short-term fluctuations. Controlling for industry-year dummies also captures industry-level shocks to skilled labor use and productivity evolution.

4.3 Individual regressions

Firm-level regressions are less suitable for estimating individual outcomes, including wage effects, therefore, we turn to individual regressions to study changes in the wage premium. Our primary interest in these regressions is the interaction of individual-level skills and the firm-level innovation status. In particular, we estimate the following equation:

$$\begin{aligned} \ln wage_{ijt} = \beta_u * innovation_{jt} + \beta_s * innovation_{jt} \times college_{it} + \delta * skillratio_{jt} + \\ + \gamma * X_{ijt} + \varphi_j + \varsigma_{st} + \varepsilon_{ijt} \end{aligned} \tag{6}$$

where i indexes employees, j firms, t years and s skill levels. $innovation_{jt}$ is the dummy showing whether the firm innovated in the corresponding CIS wave or in any of the previous two waves, and $college_{it}$ is a dummy. X_{ijt} are the usual Mincer-type controls, including gender, age, tenure, tenure squared, hours worked, a dummy for part-time employees and a dummy for new entrants. φ_j are firm

[15] We Winsorize this and the other long difference variables at the 5th and 95th percentiles.

[16] We estimate the TFP change by the ACF procedure, which already takes into account the change in inputs.

[17] The regression sample includes only firm-years when the long differences can be observed and when the bi-annual CIS was conducted. Therefore, our observations are from 2004, 2006 or 2008.

fixed effects, while ς_{st} are skill-year fixed effects, for the 4 categories of schooling.[18] We cluster the standard errors at the firm level, where the innovation status is measured.

Our main parameter of interest is β_s, or the extra college premium of innovating firms. In line with Equation (1) in our theoretical framework we also control for the skill ratio. As a result a positive β_s reflect that innovation activities are skill biased.

There are a number of econometric issues when estimating Equation (6). One important issue is the self-selection of workers based on unobservables. Higher productivity workers may be more likely to work for (or move to) innovative firms. Unfortunately, we do not have individual identifiers in the Structure of Earnings Survey (SES) that would allow us to include individual fixed effects in the regression, but we observe whether the worker is new entrant or incumbent. As a result, we deal with the selection problem in three different ways.

First, we show that our results hold if we restrict our attention to incumbent workers who were at the firm at least for 24 months. Focusing on incumbent workers ensures that our results are not driven by the higher wages of the new entrants.

Second, we follow Koren & Csillag (2017), working on the same data, in creating a non-parametric proxy for unobserved skills in the following way. First, we rank the workers in every occupation-year cell by their wages. p_{iot} denotes the wage percentile of worker i in occupation o in year t. Then we regress p_{iot} on gender, highest finished education level and birth year in every year, and denote the fitted value of this regression by $\hat{p_i}$, which we include into our regressions. The higher value of $\hat{p_i}$ means that the worker have a higher expected wage while it is not correlated with the firm characteristics.

Third, by exploiting the survey design of the SES we can (probabilistically) link observations when a worker with the same date of birth, gender, education level and occupation is observed multiple times in a firm. In the Table A9, we show that the results are robust when including individual fixed effects in the subset of workers where we are able to create this quasi panel.

Another econometric issue with the regression above is firm heterogeneity in terms of their wage levels. It is easily possible, that 'better' firms (in terms of, say, productivity or management capabilities) are both more likely to innovate and pay higher wages. Firm fixed effects are included to handle this issue.

Moreover, even after controlling for firm fixed effects, it is still possible that unobserved firm characteristics are correlated with the skill premium. For example, firms with a more decentralized organization may pay both a higher college premium and may be more likely to introduce innovations. This would imply a pre-trend in terms of the college premium. Indeed, a clear pre-trend shows up in the fixed effects estimates. We handle this problem by propensity score matching. The idea here is to focus on firms which are not innovative in the first CIS wave in which we observe them. A subset of these firms starts to innovate sometime in the future, and our aim is to create a control group for

[18]We find it especially important to control for skill-year effects because these capture skill-group-level wage trends (Acemoglu & Autor 2011, DiNardo et al. 1996) and many policy changes, including changes in the minimum wage and the expansion of higher education.

them from firms which do not.[19]

The steps of the matching procedure are the following. First, we run a probit regression with the innovation dummy as the dependent variable and basic firm characteristics as explanatory variables, with restricting the sample to each firm's first record in the CIS. The explanatory variables include both balance sheet information and a number of variables from the CIS, as suggested by Griffith et al. (2006). The latter characterize the main market of the firm, the types of funding it received and its main information sources.[20] Based on this probit, we estimate a propensity score to innovate for each firm. Second, we restrict our sample to firms which were sampled at least twice in the CIS, and were not innovative in the first period. We consider the firms which started to innovate sometime later as treated. We use propensity score matching[21] to design a control group for these firms from those which did not innovate in any of the subsequent periods, and use this sample and the resulting weights as our matched sample.

With this information-rich matching strategy we are likely to be able to focus on innovators and non-innovators which are on a common support, by excluding frequent innovators and firms which are very unlikely to innovate. This presumption is reinforced by the fact that no pre-trend is detectable in this sample. As an alternative strategy, we restrict the sample to switcher firms, and show that the results are robust to identifying the relationship only from timing. We think that it is reasonable to assume that our estimates capture at least a partly causal relationship between innovation and the higher premium. Note that, as opposed to the case of R&D, one is unlikely to find an instrument which would affect innovation strongly, but would not have an impact on the wage structure via other channels.

An additional problem is measurement error. A relevant question is the extent to which innovation is measured precisely from such a questionnaire. Undoubtedly, such surveys generate measurement errors, though most likely it works against finding significant effects. However, a number of observations convince us that the innovation variable reflects meaningful, if noisy variation.[22] First, descriptive results in, for example, Table 1 show that the innovation variable is correlated in the expected way with other firm-level characteristics. Second, our matching strategy is likely to handle this issue by excluding firms which are not on the common support, and may report 'no innovation' by mistake. We also instrument innovation with the instruments from the CIS suggested by Griffith et al. (2006), which are probably less likely to be mismeasured.

[19] An alternative is to estimate firm-skill fixed effects. This, however, has relatively little power. Still, it yields positive, though somewhat smaller estimates for the extra collage premium of the innovative firms. Alternatively, we have run regressions with interactions of propensity score deciles and the college dummy, which yield similar estimates to our preferred specification.

[20] The variables from the balance sheets are: 1-digit industry dummies, year dummies, log employment, log productivity, log wage premium, ownership. The dummies from the CIS indicate whether the worker's firm's main market is international, whether it received funding from local government, the national government, or the EU, and whether international sources, buyers, suppliers, competitors, universities or conferences were important information sources. The main results are not sensitive to using other sets of variables, for example, to excluding the CIS variables from the matching.

[21] Our main specification is a 1-nearest neighbor matching, and we report robustness tests with kernel matching. Other matching procedures yield similar results.

[22] A potential concern with the diff-in-diff estimator is that some always or frequent innovators report no innovation in the first period, and we attribute some of their premium to a switch in innovation status.

The fourth problem is the already mentioned issue with timing, which prevents us from conducting a sharp event study analysis. We handle this issue by using a relatively long window to measure the effects of innovations which captures the medium-term effects on wages.

We find that the matching strategy is a clear an intuitive way to handle these econometric problems, therefore we chose that as our preferred specification and report a number of robustness checks to validate our estimates.

4.4 Testing for heterogeneity

The first question we ask about the heterogeneity of innovation is whether the effect of the innovation depends on its novelty value, as defined in 3.1. We test for this by including the interaction of the key variables with the novelty dummies into Equation (6):

$$\begin{aligned} \ln wage_{ijt} = \beta_u * innovation_{jt} + \beta_s * innovation_{jt} \times college_{it} + \\ + \beta_u^n * innovation_{jt} \times novelty_{jt} + \beta_s^n * innovation_{jt} \times college_{it} \times novelty_{jt} + \\ + \gamma * X_{ijt} + \varphi_j + \varsigma_{st} + \varepsilon_{ijt} \end{aligned} \quad (7)$$

Our main parameter of interest is still β_s, which shows the effect of low-novelty innovation on the college premium in this setting.[23] β_s^n, the coefficient of the triple interaction term, is also of interest, because it shows whether the college premium differs between firms conducting low- and high-novelty innovation.

The second type of heterogeneity we test for is between different types of innovation, most importantly between technological and organizational innovation. Note that these innovation types are not mutually exclusive: a firm can conduct both technological and organizational innovation. Therefore, we introduce separate dummies for the different types of innovation and their interaction with the college dummy into Equation (6).[24]

4.5 Decomposition of the aggregate skill premium

The results of regression (7) can be used for a very simple decomposition of the cross-sectional aggregate skill premium to the contribution of high- and low-novelty innovators.

We do so in three steps. First, we calculate the observed wage premium for our regression sample by comparing the average log wage of college educated and other workers.

[23] Note that the novelty dummies are only defined for innovative firms, therefore, there are three types: non-innovators, low-novelty innovators and high-novelty innovators, all captured by this specification.

[24] Note that this specification assumes additivity of the effect of different innovation types. We have run regressions to test for it, and did not find any evidence that they should be treated differently.

Next, we quantify the effect of high-novelty innovation by relying on a counterfactual scenario in which high-novelty innovators conduct only low-novelty innovation. We do so by switching the R&D dummy to zero for all firms, and predicting the wages from Equation (7) for all workers. We can calculate the counterfactual wage premium by comparing the predicted wages of college educated and other workers. The difference between this wage premium and the observed wage premium shows the contribution of high-novelty innovation to the aggregate wage premium.

In the second counterfactual exercise, we attempt to quantify the contribution of all innovations on the wage premium. We use a similar strategy to the previous one, but switch the innovation variable, rather than the R&D dummy, to zero for all firms. Again, we predict wages from Equation (7).

Naturally, this exercise provides a partial way to capture the contribution of different types of innovation to the wage premium or SBTC. A key omitted mechanism is the potential reallocation of workers between innovative and non-innovative firms. Even so, we find it a useful way to check whether the mechanisms we study are likely to have an aggregate relevance.

5 Results

5.1 Country-industry-level evidence

Figure A2 illustrates the industry-country level relationship between broadly defined innovation according to the 2010 CIS (which captures innovative activities between 2008 and 2010) and subsequent growth in skill demand. The figures suggest a clear positive relationship between innovation and both the quantity and the wage response.

Table 3 presents the regression results both for the change in the share of college educated workers (upper part) and their wage premium (lower part). Column (1) reports basic regressions when both the share of innovative firms and the R&D intensity are included.[25] The estimates suggest that the increase in skill demand is linked to broadly defined innovation rather than only R&D. A 10 percentage point higher share of innovative firms is associated with 1 percentage points stronger growth of the college employment share and 3 percentage points higher increase in the college premium at the industry level. The R&D variable is small and often has a negative point estimate.

Column (2) includes country fixed effects to control for country-level shocks in skill supply or economics growth, column (3) includes industry fixed effects while column (4) includes both. In the college share regressions, the results remain unchanged when industry fixed effects are included, while the innovation coefficients become insignificant industry fixed effects are included. The point estimates in the college premium equations are similar independently of the types of fixed effects included, suggesting a strong relationship between innovation and subsequent increase in the college

[25]Including only the broadly defined innovation measure yields similar results or controling for high-novelty innovation with the share of R&D-conducting firms in the CIS yields similar results.

premium. In Table A4, in line with our individual specifications, we also control for the change in the share of college educated workers, and find that the results are robust for this extension.

We can conclude from this exercise that broadly defined innovation is more strongly related to increasing skill demand than R&D. Low-novelty innovation, including technology adoption can be as significant in SBTC as high novelty innovation. Also, we see a response both at the quantity and the wage margin at the industry level, which motivates our investigations of both margins at the firm level.

5.2 Firm-level evidence

Table 1 compares innovative and non-innovative firms in our sample. Two types of differences are apparent. First, in line with much of the literature (Griffith et al. 2006), innovative firms are larger, more capital intensive and more productive. Second, innovation is indeed associated with higher skill levels and higher wages. In particular, innovative firms have 5 percentage point more college graduates and pay about 20 percent higher wages, on average.

Table 2 shows the role of between-firm variation in total wage variation, following the decomposition of Song et al. (2015). Firm fixed effects explain about 50% of wage variation, which is relatively large in international comparison.[26] The innovation dummy explains about 4.9 percentage points from this. Importantly, the explanatory power of the innovation dummy is about 50% higher than that of the R&D dummy. Including the different types of innovation (R&D, product, process, organization) into the regression increases the explanatory power further, to 7.9 percent. Another apparent pattern is that, in line with the more heterogeneous nature of high-skilled labor, both firm effects and innovation explain more from low-skilled wage variation than from the the wage differences of the highly skilled.

The firm-level regression results are presented in Table 4. We start with modeling the long difference of the wage share of college educated workers in column (1). We find a significant positive relationship, suggesting skill upgrading in innovative firms, where the wage share increased by 1.7 percentage points during a 6-year period. Column (2) shows that this was, at least partly, explained by the increasing employment share of these worker. According to column (3), the wage effects can only be estimated imprecisely, which is one of the motivations of using worker-level regressions to estimate this.[27] Columns (4) and (5) confirm, by using two measures of TFP, that innovation is associated with stronger subsequent productivity growth, in the order of 1 percentage point per year. Finally, column (6) shows no evidence for stronger size growth of innovative firms.

A potential concern with the long difference specifications is that their results may be sensitive to the choice of the length of the difference. In Table A5 we re-run the main regressions with 2, 4 and 6-year long lags to find a positive association between innovation and subsequent growth both in

[26] Song et al. (2015) report that between 2007-2013, between-firm inequality explained 42.1 percent the variance of wages.

[27] One reason for this is that we only observe the dependent variable, the 6-year change in the ratio of the the wage of college educated and other workers for relatively few firms.

college share and TFP. The effects are stronger for longer differences suggesting that innovation has prolonged effects.

5.3 Worker-level evidence

The main individual-level results are presented in Table 5, when the college share is included in the regression. Table A6 presents the results without this variable. Column (1) shows results on the full sample, when only skill-year fixed effects are included. According to these results, workers without a college degree earn 14.6 percent more in innovative firms (relative to workers with similar education levels in non-innovative firms), while this difference is 26.7 percent for college educated workers (compared to collage educated workers in non-innovative firms). Importantly, the skill ratio is positively associated with wages: firms relying on more skill-intensive technologies pay higher wages for their workers in general.

In column (2) we also control for worker observables. In general, controlling for worker observables changes only slightly the innovative firms' wage premium. Column (3) also includes the $\hat{p}$ variables to control for worker unobservables. Again, though highly significant, innovators' premia remain very similar.

We also include firm fixed effects in column (4). In this specification, the low-skilled innovation premium disappears, while the college innovation premium becomes even higher than before, at 11% relative to college educated workers of non-innovative firms. This suggests that while innovative firms pay higher wages even before the innovation, the innovation itself is associated only with an increase in the wages of the high-skilled.

Importantly, the sign of the skill ratio becomes negative and significant when we control for unobserved firm heterogeneity. This is strongly in line with Equation (1) . The estimated coefficient, between -0.12 and -0.17 suggests that the firm-level elasticity of substitution is between -5.8 and -8.3.

In column (5), we also include a pre-trend variable, which takes the value of 1 if the firm does not innovate in a CIS wave, but will innovate in the next wave. This pre-trend is positive and significant for college educated workers, suggesting that firms paying a higher college premium are indeed more likely to innovate subsequently. Still, even according to this specification, the college premium increases by about 4.5 percentage points following the innovation.

The presence of this pre-trend is a motivation for the matching strategy, our preferred specification, the results of which are presented in columns (6) and (7).[28] Clearly, there is no pre-trend in this sample. Consistently with the previous results, we find that innovation is only associated with an increase in the wage of college educated workers. Following the innovation, the college premium increases significantly, by about 8.3 percent.

A number of robustness checks are presented in Table 6. First, column (1) includes (1-digit)

[28] Importantly, the results are similar if we do not weight by the matching weights but only restrict the sample to the common support, or consider all firms which were not innovative in the first period.

industry-skill-year fixed effects to check whether the college premia results from an industry composition effect. Similarly, in column (2) we include (2-digit) occupation-year fixed effects.[29] In column (3), we define the innovation variable based only on the corresponding CIS wave. Table A7 presents these specifications in more detail and confirms that there is no pre-trend under these alternative definitions. Fourth, in column (4), we rely on kernel matching rather than the nearest neighbor matching used in the main model. In column (5), we restrict the sample to switcher firms, which were initially non-innovative but started to innovate at some point.[30] In column (6) of Table 6 instruments the innovation variable with information from the CIS, as proposed by Crépon et al. (1998) and Griffith et al. (2006).[31] While these variables may not be exogenous, they, taken together, may be less prone to measurement error than the innovation variable. Finally, in column (7) we exclude foreign-owned firms, which may be more innovative and have a different wage structure than domestically-owned ones. All in all, the main results are robust to all these robustness checks.

As we have discussed, a key concern with the data at hand is that, even when controling for worker observables, the results can be driven by changes in the composition of workers. Table A8 restricts the sample to such workers, defined as had been working for the firm for at least 24 months. We find that results on this subsample are very similar to the full sample, suggesting that incumbent college educated workers, and not only new hires, experienced an increase in their premium.

Another way of tackling this issue is to create a quasi-panel from the data by linking the workers probabilistically. Unfortunately only a minority of workers can be identified in such a way, and these workers tend to work in large firms, which are sampled in every year. Worker fixed effects can only be applied on this subsample. In Table A9 we restrict the sample to workers whose wage is used in a within-worker identification strategy, i.e. who are observed multiple times and who worked at the firm both before and after an innovation. Columns (1) and (2) show that the observed extra college premium of innovators is 4.4 percentage on this sample independently of whether firm fixed effects are included. In column (3) we include $\hat{p}$ and restrict the sample to incumbent workers, to replicate the results of Table A8 on this smaller sample. The estimated extra college premium resulting from innovation declines to 3 percent. Finally, in column (4) we include worker fixed effects to find a significant (at 5%) premium of 2.3 percent.

One can draw two lessons from this exercise. First, innovation is associated with a higher college premium even when worker fixed effects are included. This estimated premium may be biased downward, because it is estimated on a selected sample where the OLS college premium is also much lower than on the full or the matched sample. Furthermore, workers in the quasi-panel are identified partly based on their occupation; therefore, we cannot follow them once they are promoted, which may be a dominant way to raise wages. Second, results with and without worker fixed effects are very similar in this sample; therefore, one can hope that worker fixed effects would not matter that much on the full sample after one includes firm fixed effects, $\hat{p}$ and restricts the sample to incumbents.

[29]Importantly, this controls for the potentially higher wages paid to R&D staff. Further, if we exclude workers with an R&D occupation, the results remain similar.

[30]The results are robust if we only consider switchers which are part of the matched sample.

[31]These variables indicate whether the worker's firm's main market is international, whether it received funding from local government, the national government, or the EU, and whether international sources, buyers, suppliers, competitors, universities or conferences were important information sources. The other variables used by Griffith et al. (2006) were not available in all years of the CIS.

5.4 Heterogeneity of innovation

Let us start investigating heterogeneity with firm-level regressions, reported in Table 7. In the odd-numbered columns, we test for the importance of novelty of innovation by including the interaction of innovation and the R&D dummy, as in Equation (7). Now the coefficient of innovation captures the effect of low-novelty innovation while the interaction captures the difference between low- and high-novelty innovation. We find some evidence suggesting that low-novelty innovation is related both to an increase in the college employment share and to an increase in productivity growth. R&D-based innovation is even more strongly related to these outcomes, especially to the magnitude of productivity growth. In even-numbered columns, we attempt to distinguish between technological and organizational innovation. We find that it is not easy to disentangle these two types of innovation in this specification, though organizational innovation remains significantly correlated with college share growth.

Table 8 estimates Equation (7) with the three novelty dummies: (i) R&D; (ii) whether the product/process was new to the market; (iii) whether the product process was developed by the firm or adopted. The overall picture is that low-novelty innovations are associated with a substantial extra college premium, whichever way one controls for novelty. Importantly, according to these regressions, low- and high-novelty innovation is indistinguishable in terms of their effect on the firm-level college premium. Both appear to be similarly skill-biased.

Table 9 distinguishes between different types of innovations. These regressions suggest that technological innovation is associated with a higher college premium than organizational innovation. Distinguishing between product and process innovation is possible only less precisely. That said, the coefficients of process innovation are slightly larger than that of product innovation.

Altogether, these results are in line with the hypotheses that low-novelty innovation is skill-biased, and that the magnitude of its bias is similar to that of high-novelty innovation. Further, distinguishing between different types of low-novelty innovations, we find that technological innovation - probably mainly due to its process innovation sub-component - may be more skill-biased than organizational innovation.

5.5 Mechanisms

In this subsection, we present a few pieces evidence on the mechanisms involved. Note that some of the earlier results have provided us with useful information about how SBTC takes place. As Table A7 has shown, the estimated effect does not depend much on the time period used. This suggests that the skill premium is long-lasting, and not constrained to the period when the innovation takes place. Table A10 complements this evidence with regressions on the matched sample where we replace the the dependent variable with other worker-level outcomes. A comparison of columns (1) and (2) shows that the increase in the base wage after the innovation was similar to the increase in the total wage of the worker. In addition, column (3) documents that the probability of receiving any bonus payments does not change when the firm conducts innovations. Finally, column (4) shows that the increase in the skill premium does not result from an increase in the number of hours worked. Taken together, these pieces of evidence are likely to reflect long-run changes in the operations of the firm rather than temporary bonuses for the increased effort accompanying the innovation process itself.

A frequently cited mechanism is that technological change affects differently workers performing routine and non-routine tasks Autor et al. (2003). A college degree may be strongly correlated with non-routine occupations, and increasing college premium may capture this aspect of work. To investigate this possibility, in Table A11 we include the measure proposed by Autor et al. (2003) and its interaction with innovation besides college and its interaction.[32] We include $1 - RTI$ so that it is increasing in non-routine content. We find that people working in less routine jobs are paid higher wages in general. With the exception of the matching specification, innovative firms pay a higher premium for workers with non-routine jobs. This provides evidence for technological change which is biased against occupations requiring more routine tasks. However, including these variables does not affect substantially the estimates for the college premium when firm fixed effects are included: innovation seems to favor college educated workers even when we control for the routine task content of their jobs.

A final question is whether the increased college premium is part of a picture where only the wages of highly qualified workers increase or is it a symptom of wage polarization. Table A12 reports results when all four skill levels are interacted with the innovation variable. Note that the omitted category is those with secondary schooling, and the interactions show the change in wages following an innovation relative to this category (again, after controlling for skill-year fixed effects). The results in the table provide little evidence for polarization, neither in cross section nor for firm fixed effects specifications. In actual fact, the wages of the lower three educational categories do not seem to change after innovation takes place, while the wages of college educated workers do increase substantially.

5.6 Quantifying the importance of innovators in the aggregate skill premium

Panel A of Table 10 presents the results of the simple decomposition exercise described in Section 4.5 on the matched sample. According to column (1), taking the observed wages of all workers, the wage premium of college educated workers is 80 (log) percent.

Column (2) shows the results of our first counterfactual exercise, which aims at quantifying the role of high-novelty innovation. Accordingly, we 'switch' all high-novelty innovators to low-novelty innovators and predict workers wages based on the regression presented in column (2) of Table 8. We find that the college premium would decrease by only 0.3 percentage points.

Column (3) presents the results of the second counterfactual exercise, which attempts to quantify the role of low-novelty innovation. Indeed, if no firm had innovated, the college premium would be 74.3 percent, or 5.8 percentage points lower than what is actually observed in the data. This provides evidence that low-novelty innovation can have strong aggregate effects on the college premium and

[32] We link the US occupation codes to Hungarian occupation codes.

can be a forceful driver of SBTC at the level of the economy.

One potential concern with this exercise is that the matched sample may not be representative for the economy in general. Most importantly, few firms are likely to switch from non innovation to high-novelty innovation. Therefore, as a robustness check, we repeat this exercise on the full sample. We also re-estimate equation (7) on this sample. While the role of high-novelty innovation is somewhat larger in this sample (1.7 percentage points), that of low-novelty innovation remains dominant (6.2 percentage points).

6 Conclusions

This paper uses a very rich dataset which combines detailed firm-level information on innovation with worker-level wage data from Hungary. Based on the innovation survey, we rely on a very broad definition of innovation, which includes product, process and organizational innovation independently of its novelty value. Using panel and matching identification strategies, we find that innovation defined in such a way is positively associated both with an increase in the share of college educated workers and their wage premium. We also show that the novelty value of the innovation, respective of how it is measured, is not strongly associated with the skill premium: indeed, high- and low-novelty innovation seem to be similarly skill-biased in terms of the wage premium. We also find that, in quantitative terms, low-novelty innovation contributes strongly to the aggregate college premium.

The key conclusion from these results is that skill-biased technological change is not necessarily linked to generating new knowledge or high novelty products at the firm level. This finding does not contradict influential theories of SBTC. On the contrary, those theories often emphasize technology diffusion or relatively low-novelty follow-up innovations as key sources of economy-wide technological change.

From a theoretical point of view, our results that technological, and especially process, innovations are more strongly associated with the skill premium underline that technology-skill complementarity may be a key mechanism behind SBTC. The fact that product and organizational innovation are also associated with an increase in the skill premium suggests that other mechanisms, including the skill bias resulting from organizational change, may also be at play.

For policymakers, the main message of these results is that skill-biased technological change and the resulting inequality may be affected by more factors than traditional R&D activities (OECD 2015). Competition, globalization or access to different types of knowledge may drive such technological change, which should all be taken into account when evaluating different policy alternatives. In a more globalized world, stronger knowledge flows lead to more technology adoption and low-novelty innovation, and, therefore, an increased skill premium (Keller 2004).

References

Abowd, J. M., Haltiwanger, J., Lane, J., McKinney, K. L. & Sandusky, K. (2007), Technology and the demand for skill: An analysis of within and between firm differences, Technical report, National Bureau of Economic Research.

Acemoglu, D. (2002), 'Technical change, inequality, and the labor market', Journal of Economic Literature **40**(1), 7–72.

Acemoglu, D. (2003), 'Patterns of skill premia', The Review of Economic Studies **70**(2), 199–230.

Acemoglu, D. & Autor, D. (2011), 'Skills, tasks and technologies: Implications for employment and earnings', Handbook of Labor Economics **4**, 1043–1171.

Ackerberg, D. A., Caves, K. & Frazer, G. (2015), 'Identification properties of recent production function estimators', Econometrica **83**(6), 2411–2451.

Aghion, P., Bergeaud, A., Blundell, R. & Griffith, R. (2017), 'Innovation, firms and wage inequality'.

Akerman, A., Gaarder, I. & Mogstad, M. (2015), 'The skill complementarity of broadband internet', The Quarterly Journal of Economics **130**(4), 1781–1824.

Akhmetova, Z. & Ferguson, S. M. (2015), 'Offshoring, total factor productivity and skill-biased technological change'.

Autor, D. H., Katz, L. F. & Krueger, A. B. (1998), 'Computing inequality: have computers changed the labor market?', The Quarterly Journal of Economics **113**(4), 1169–1213.

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.

Bayoumi, T., Coe, D. T. & Helpman, E. (1999), 'R&d spillovers and global growth', Journal of International Economics **47**(2), 399–428.

Bloom, N., Sadun, R. & Van Reenen, J. (2016), Management as a technology?, Technical report, National Bureau of Economic Research.

Bøler, E. A. (2015), 'Technology-skill complementarity in a globalized world', University of Oslo Mimeo .

Bresnahan, T. F., Brynjolfsson, E. & Hitt, L. M. (2002), 'Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence', The Quarterly Journal of Economics **117**(1), 339–376.

Bustos, P. (2011a), 'The impact of trade liberalization on skill upgrading', Evidence from Argentina .

Bustos, P. (2011b), 'Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms', American Economic Review **101**(1), 304–40.

Card, D., Cardoso, A. R., Heining, J. & Kline, P. (2018), 'Firms and labor market inequality: Evidence and some theory', Journal of Labor Economics **36**(S1), S13–S70.

Caroli, E. & Van Reenen, J. (2001), 'Skill-biased organizational change? evidence from a panel of british and french establishments', The Quarterly Journal of Economics **116**(4), 1449–1492.

Caselli, M. (2014), 'Trade, skill-biased technical change and wages in mexican manufacturing', Applied Economics **46**(3), 336–348.

Coe, D. T. & Helpman, E. (1995), 'International r&d spillovers', European Economic Review **39**(5), 859–887.

Costantini, J. & Melitz, M. (2008), The dynamics of firm-level adjustment to trade liberalization, in E. Helpman, D. Marin & T. Verdier, eds, 'The organization of firms in a global economy', Harvard University Press Cambridge, MA, chapter 4, pp. 107–141.

Crépon, B., Duguet, E. & Mairessec, J. (1998), 'Research, innovation and productivity: An econometric analysis at the firm level', Economics of Innovation and New Technology **7**(2), 115–158.

DiNardo, J., Fortin, N. M. & Lemieux, T. (1996), 'Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach', Econometrica **64**(5), 1001–1044.

Fagerberg, J. (2007), 'A guide to schumpeter', Confluence: Interdisciplinary Communications **2008**, 20–22.

Goldin, C. & Katz, L. F. (2010), The race between education and technology, Technical report, Belknap Press.

Griffith, R., Huergo, E., Mairesse, J. & Peters, B. (2006), 'Innovation and productivity across four european countries', Oxford Review of Economic Policy **22**(4), 483–498.

Kasahara, H., Liang, Y. & Rodrigue, J. (2016), 'Does importing intermediates increase the demand for skilled workers? plant-level evidence from indonesia', Journal of International Economics **102**, 242–261.

Keller, W. (2004), 'International technology diffusion', Journal of Economic Literature **42**(3), 752–782.

Kline, P., Petkova, N., Williams, H. & Zidar, O. (2018), Who profits from patents? rent-sharing at innovative firms, Technical report, National Bureau of Economic Research.

Koren, M. & Csillag, M. (2017), Machines and machinists: Importing skill-biased technology, Technical report, Mimeo, Central European University.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', The Review of Economic Studies **70**(2), 317–341.

Machin, S. & Van Reenen, J. (1998), 'Technology and changes in skill structure: evidence from seven oecd countries', The Quarterly Journal of Economics **113**(4), 1215–1244.

Mairesse, J. & Mohnen, P. (2010), Using innovation surveys for econometric analysis, in B. H. Hall & R. Nathan, eds, 'Handbook of the Economics of Innovation', Vol. 2, Elsevier, pp. 1129–1155.

Mokyr, J. (2003), 'Thinking about technology and institutions', Macalester International **13**(1), 33–66.

OECD (2015), The innovation imperative: Contributing to productivity, growth and well-being, OECD Publishing.

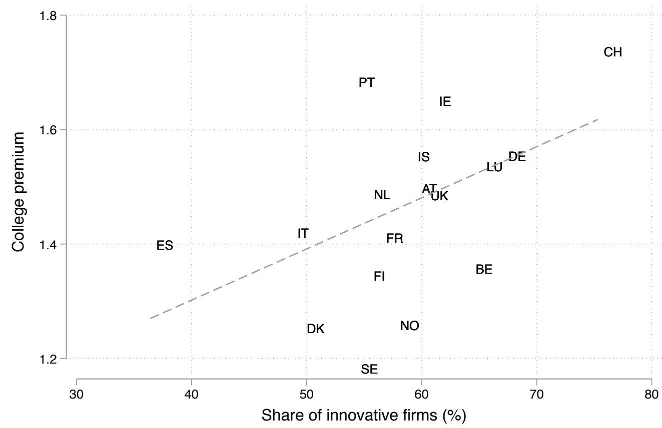
Song, J., Price, D. J., Guvenen, F., Bloom, N. & Von Wachter, T. (2015), Firming up inequality, Technical report, National Bureau of Economic Research.

Verhoogen, E. A. (2008), 'Trade, quality upgrading, and wage inequality in the mexican manufacturing sector', The Quarterly Journal of Economics **123**(2), 489–530.

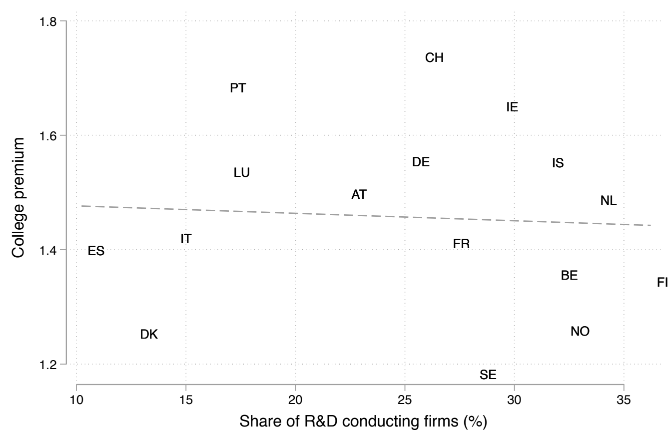
Figures

Figure 1: Innovation and the college premium: cross-country evidence

Panel A: Innovative firms



Panel B: R&D conducting firms



Notes: This figure shows the cross-country relationship between the college premium in 2014 and the share of innovative firms (Panel A) and the share of R&D conducting firms (panel B), according to the 2014 Community Innovation Survey for old EU member states. Innovative firms are firms conducting broadly defined innovation activities, including product, process, organizational or marketing innovations which may or may not involve R&D activities.

Tables

Table 1: Comparing innovative and non-innovative firms

Variable	Non-innovative	Innovative	diff	t-value
Average age of empl.	42.1	41.3	-0.8	-6.94
	(0.09)	(0.10)		
Share of female empl.	0.21	0.19	-0.02	-3.43
	(0.01)	(0.00)		
Average year of education	11.4	11.8	0.3	11.09
	(0.02)	(0.03)		
Share of college grad.	0.12	0.18	0.05	12.73
	(0.00)	(0.00)		
Average wage	173,087	206,746	33,659	12.90
	(1,672.11)	(2,446.14)		
Foreign-owned (dummy)	0.31	0.41	0.11	8.71
	(0.01)	(0.01)		
Number of employees	159	435	276	6.91
	(7.43)	(44.85)		
ln(tangible capital/employees)	7.95	8.46	0.51	15.08
	(0.03)	(0.03)		
ln(value added/employees)	8.24	8.54	0.30	14.72
	(0.01)	(0.02)		

Note: This table compares innovative and non-innovative firms in terms of key variables and tests whether the difference between the two groups is significant. One observation in this table is one firm-year and the sample includes all firms which were sampled by the Community Innovation Survey between 2004 and 2014. Innovative firms are those which conduct product, process or organization innovation according to the Community Innovation Survey (CIS). The second and third columns show the average value of the variable for the two groups of firms with its standard deviation in parentheses below. The last two columns show the difference between innovative and non-innovative firms and the corresponding t-statistic. The source of variables in the different rows is the balance sheet data, where nominal variables were deflated with industry-level deflators.

Table 2: The explanatory power of the innovation activities variables in wage inequality

Share of wage variation explained by:	All workers	No college	College
Firm FE	52.5%	54.7%	44.2%
R&D Dummy	3.3%	3.2%	1.5%
Innovation dummy	4.9%	4.5%	2.3%
Innovation dummy + R&D dummy	5.2%	4.8%	2.5%
Type of innovation dummies	7.9%	7.6%	3.3%

Note: This table shows the share of wage variation explained by different innovation dummies. In particular, it reports the R-squared of cross sectional regressions for 2014 with log wage as dependent and (i) firm fixed effects, (ii) an R&D dummy, (iii) an innovation dummy, (iv) an R&D and an innovation dummy and (iv) innovation type dummies as explanatory variables. In the second column, the regressions were run on the sample of all workers, while in the third and fourth columns on the subsample of workers without and with a college degree, respectively. For example, the second column shows that 52.5 percent of total wage variation is explained by between-firm differences. 3.3 percentage points of this is explained by an R&D dummy, while 4.9 percentage points, or 50% more, is explained by the broader innovation dummy.

Table 3: The share of innovative firms and growth in skill demand, country-industry-level evidence

	College share change, 2010-2014			
	(1)	(2)	(3)	(4)
Share of innovative firms (2010)	0.104***	0.075	0.122***	0.011
	(0.025)	(0.049)	(0.031)	(0.050)
R&D-intensity (2010)	-0.008***	-0.000	-0.012***	-0.003
	(0.003)	(0.002)	(0.004)	(0.002)
country FE		yes		yes
industry FE			yes	yes
Observations	158	156	157	155
R-squared	0.154	0.697	0.255	0.770

	College premium change, 2010-2014			
	(1)	(2)	(3)	(4)
Share of innovative firms (2010)	0.284**	0.250**	0.185	0.242*
	(0.128)	(0.119)	(0.124)	(0.136)
R&D-intensity (2010)	-0.020**	-0.003	-0.028**	-0.007
	(0.009)	(0.006)	(0.011)	(0.006)
country FE		yes		yes
industry FE			yes	yes
Observations	154	152	153	151
R-squared	0.192	0.670	0.303	0.714

Note: These tables shows regressions at the 1-digit industry-country level for 25 European countries. The dependent variable is the change in the share of college educated workers and their skill premium. The main explanatory variable shows the share of innovative firms according to the 2010 CIS wave, measuring innovation activities between 2008 and 2010. Observations are weighted with the number of firms in the country-industry cell from the CIS. Standard errors, clustered at the country level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: Innovation and subsequent change in firm-level outcomes

LHS:	(1) college wage share	(2) college employment share	(3) wage rate	(4) TFP (ACF)	(5) TFP (LP)	(6) ln employment
Innovation	0.017***	0.019**	0.030	0.061**	0.050**	0.030
	(0.004)	(0.008)	(0.023)	(0.025)	(0.022)	(0.020)
ln capital (d)	-0.006	-0.007	-0.011			
	(0.004)	(0.007)	(0.019)			
ln value added (d)	-0.005	-0.007	0.010			
	(0.005)	(0.008)	(0.022)			
Dependent variable (t-1)	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes
Observations	2,153	2,153	1,386	2,122	2,122	2,122
R-squared	0.099	0.095	0.207	0.140	0.148	0.144

Note: This table shows the firm-level relationship between innovation and subsequent 6-year change of key variables, following Caroli & Van Reenen (2001) and specified in Equation (5). The dependent variables are the long differences of the variables in the column headings, defined as their change between t and $t+6$. The Innovation dummy shows whether the firm conducted product, process or organizational innovation between years $t-5$ and t, according to the CIS waves conducted in years t and $t-2$. The other two explanatory variables in columns (1)-(3) are long differences of log capital stock and log value added. The sample includes firms which were surveyed either in the 2004, 2006 or 2008 CIS waves, when the 6-year change can be observed. The dependent variables in columns (1)-(3) are calculated from the Structure of Earning Survey, while those in Columns (4)-(6) are calculated from Balance Sheet data. Columns (1) and (2) show the relationship between innovation and the subsequent change in the share of college educated workers in the wage bill and employment, respectively. In column (3), the dependent variable measures the change in the ratio of the average wage of college educated workers and of workers with lower levels of education. Note that this can only be observed when both college and non-college educated workers are observed for the firm in both t and $t+6$. Columns (4) and (5) investigate the relationship between innovation and TFP change, where TFP is estimated with the methods proposed by Ackerberg et al. (2015) and Levinsohn & Petrin (2003), respectively. In column (6), the dependent variable is change in employment. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 5: Innovation and the college premium: worker-level regressions

LHS: log wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Innovation	0.146***	0.112***	0.112***	-0.027**	-0.025**	-0.006	-0.008
	(0.020)	(0.016)	(0.016)	(0.013)	(0.010)	(0.009)	(0.010)
Innovation x College	0.121***	0.136***	0.123***	0.113***	0.118***	0.082***	0.083***
	(0.025)	(0.022)	(0.021)	(0.013)	(0.013)	(0.019)	(0.021)
ln(skill ratio+1)	1.096***	1.104***	1.072***	-0.123***	-0.122***	-0.171***	-0.170***
	(0.060)	(0.051)	(0.050)	(0.035)	(0.035)	(0.058)	(0.058)
Innovation pre-trend					-0.001		-0.006
					(0.020)		(0.014)
Innovation pre-trend x College					0.074***		0.011
					(0.022)		(0.023)
Skill-year FE	yes	yes	yes	yes	yes	yes	yes
Mincer variables		yes	yes	yes	yes	yes	yes
$\hat{p}$			yes	yes	yes	yes	yes
Firm FE				yes	yes	yes	yes
Matched sample						yes	yes
Observations	785,443	785,443	785,443	785,419	785,419	157,714	157,714
R-squared	0.482	0.552	0.559	0.717	0.717	0.700	0.700
Clusters	6236	6236	6236	6212	6212	1075	1075

Note: Note: This table investigates whether innovative firms pay a higher college premium with worker-level regressions, described in Equation (6), with log wage as the dependent variable. The Innovation dummy indicates whether the firm has conducted innovation either in the current CIS wave or in one of the previous two waves. The innovation x college interaction is the variable of interest, showing the extent to which the college premium is larger in innovative firms relative to non-innovative enterprises. For example, column (1) shows that non-college educated workers earn 14.6 percent more, while college educated workers earn 26.7 percent more in innovative firms relative workers with similar education levels in non-innovative firms. Skill-year fixed effects represent interactions of primary, secondary, vocational and college dummies with year dummies. Mincer variables are gender, age, tenure, tenure squared, hours worked, a dummy for part-time employees and a dummy for new entrants. $\hat{p}$ is a proxy for worker unobservables, predicted from information on occupation, age and schooling, following Koren & Csillag (2017). Our preferred specification, reported in column (6), is run on the matched sample. Here the treated group consists of initially non-innovative firms which start to innovate at some point (switchers), while the untreated group is that of never innovating firms. Propensity score nearest neighbor matching is used to find controls for treated firms from the untreated groups to handle potential pre-trends in the skill premium. Standard errors, clustered at the firm level are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 6: Innovation and the college premium: robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LHS: log wage	Industry-skill-year-FE	Occupation-year-FE	Same wave	Kernel matching	Switcher sample	CDM instruments	Only domestic
Innovation	-0.001	-0.005	-0.009	-0.008	0.004	-0.029	0.003
	(0.009)	(0.009)	(0.007)	(0.009)	(0.009)	(0.028)	(0.011)
Innovation x College	0.067***	0.080***	0.070***	0.092***	0.055***	0.081***	0.083***
	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)	(0.031)	(0.030)
Sample	matched (NN)	matched (NN)	matched (NN)	matched (kernel)	switcher	matched (NN)	matched (NN) domestic
Skill-year FE	yes	yes	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes	yes	yes
$\hat{p}$	yes	yes	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes	yes	yes
Industry-skill-year FE	yes	no	no	no	no	no	no
Occupation-year FE	no	yes	no	no	no	no	no
# waves for innov. var	3	3	1	3	3	1	3
Observations	157,709	157,714	142,249	226,422	105,143	142,249	88,661
R-squared	0.704	0.754	0.697	0.701	0.692	0.697	0.685

Note: This table shows robustness checks of the worker-level regressions, described in Equation (6), with log wage as the dependent variable. All specifications here start from the preferred specification, column (6) of Table 5. In columns (1) and (2) industry-skill-year and occupation-year fixed effects are also included, respectively. In column (3), the innovation variable is defined based only on the current CIS wave rather than the last 3, as in the base specification (see also Table A7). In column (4), kernel matching is used to define the control group rather than nearest neighbor matching. In column (5), we restrict the sample to initially non-innovative firms which switch to innovating at some point. Therefore, the coefficients are only estimated from time variation. In column (6), the innovation variable is instrumented with information from the CIS on the firms most important market, sources of funding and information sources, following Crépon et al. (1998) and Griffith et al. (2006). In column (7), the sample is restricted to domestically-owned firms. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 7: Type of innovation and subsequent change in firm-level outcomes

LHS:	College emp. share		College wage share		TFP (ACF)		TFP (LP)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Innovation	0.011**		0.007		0.035		0.030	
	(0.005)		(0.009)		(0.028)		(0.026)	
Innovation x R&D	0.013**	0.010	0.026***	0.025**	0.055*	0.067*	0.045	0.058*
	(0.006)	(0.006)	(0.010)	(0.010)	(0.033)	(0.035)	(0.029)	(0.031)
Technological inn.		0.007		0.004		-0.008		-0.018
		(0.005)		(0.009)		(0.032)		(0.030)
Organizational inn.		0.009*		0.005		0.026		0.032
		(0.005)		(0.008)		(0.029)		(0.025)
Value added (d)	yes	yes	yes	yes				
Capital (d)	yes	yes	yes	yes				
Dependent var. (t-1)	yes	yes	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes
Dependent var. (t-1)	yes	yes	yes	yes	yes	yes	yes	yes
Observations	2,153	2,153	2,153	2,153	2,122	2,122	2,122	2,122
R-squared	0.102	0.103	0.100	0.100	0.141	0.141	0.149	0.149

Note: This table extends the firm-level regressions in Table 5 to study the heterogeneity of innovation activities. Odd-numbered columns test whether high-novelty innovation, proxied by an R&D dummy, is associated with higher growth of the skilled share and TFP than low-novelty innovation. To this end, these specifications include the interaction of the innovation and R&D dummy in addition to the innovation variable. The coefficient of this interaction term shows the premium of firms conducting high-novelty innovation relative to firms conducting low-novelty innovation in terms of the 6-year change of the dependent variables. For example, according to column (1), relative to non-innovators, low-novelty (high-novelty) innovators experience 1.1 (2.4) percentage points higher increase in the employment share of college educated workers. In even-numbered columns innovation is decomposed into technological (product or process) and organizational innovation, with controlling for R&D. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 8: The novelty of innovation and the college premium

LHS: log wage	(1)	(2)	(3)	(4)	(5)
Innov x college	0.083***	0.080***	0.090***	0.082***	0.082***
	(0.020)	(0.021)	(0.020)	(0.020)	(0.022)
Innov x R&D x college		0.005			0.015
		(0.026)			(0.025)
Innov x new x college			-0.028		-0.037
			(0.031)		(0.033)
Innov x developed x college				0.004	0.007
				(0.024)	(0.021)
Innovation, novelty	yes	yes	yes	yes	yes
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes
$\hat{p}$	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes	yes
Observations	157,714	157,714	157,714	157,714	157,714
R-squared	0.699	0.699	0.699	0.699	0.699

Note: This table investigates whether high-novelty innovation is associated with a higher college premium than low-novelty innovation. To this end, it extends the worker-level regressions, reported in column (6) of table 5, by interacting the innovation x college term with different dummies proxying for the novelty of innovation. The coefficients of these triple interaction terms represent the extra college premium of high-novelty innovators relative to low-novelty innovators, whose extra college premium relative to non-innovators is shown by the coefficient of innovation x college. The columns differ in how high-novelty is measured: in column (2), it is an R&D dummy, in column (3) it is dummy showing whether the product innovation was new to the country or the process innovation was new to the market, while in column (4) it shows whether the product or process was developed by the firm rather than adopted. For example, according to column (2), relative to non-innovative firms, the college premium was 8 (8.5) percent higher in non-R&D innovators (R&D innovators). Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 9: The type of innovation and the college premium

LHS: log wage	(1)	(2)	(3)	(4)
Tech. x college	0.078***		0.087***	
	(0.020)		(0.023)	
Org x College	0.022	0.024	0.024	0.025
	(0.021)	(0.021)	(0.020)	(0.021)
Process x college		0.053**		0.058**
		(0.025)		(0.023)
Product x college		0.030		0.036
		(0.027)		(0.031)
Innov x R&D x college			-0.020	-0.018
			(0.028)	(0.030)
Innovation type, novelty	yes	yes	yes	yes
Skill-year FE	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes
$\hat{p}$	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes
Observations	157,714	157,714	157,714	157,714
R-squared	0.699	0.699	0.699	0.699

Note: This table investigates how product, process and organizational innovation are associated with the college premium. To this end, it extends the worker-level regressions, reported in column (6) of table 5 by splitting the innovation dummy into technological (product+process) and organizational innovation in columns (1) and (3) and into product, process and organizational innovation in columns (2) and (4). Columns (3) and (4) also control for novelty, proxied by the R&D dummy. Note that a firm can conduct multiple types of innovation at the same time. For example, according to column (1), relative to non-innovator firms, the college premium is 7.8 (2.2) percentage points higher at firms conducting only technological (organizational) innovation, while this extra premium is 9 percentage points at firms which conduct both technological and organizational innovation. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 10: How much does low and high innovation explain from the wage premium?

Panel A: Matched sample

ln (wage):	(1) Observed	(2) No high innov.	(3) No innov
Low skilled wage	12.130	12.129	12.138
College wage	12.930	12.926	12.880
College wage premium	0.800	0.797	0.742

Panel B: Full sample

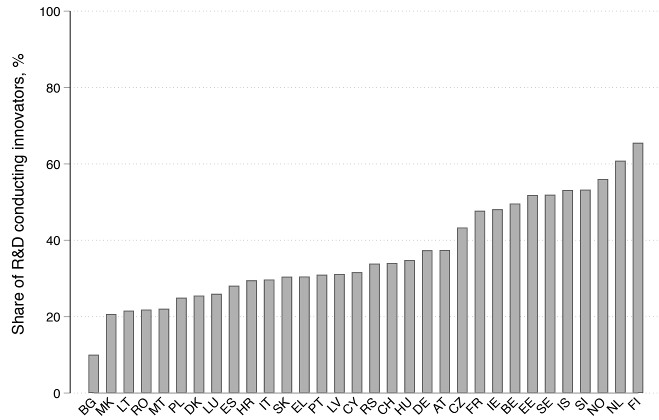
ln (wage):	(1) Observed	(2) No high innov.	(3) No innov
Low skilled wage	12.134	12.137	12.150
College wage	12.947	12.933	12.884
College wage premium	0.813	0.796	0.734

Note: This table uses a decomposition exercise to quantify the extent to which high-novelty and low-novelty innovators contribute to the cross-sectional college premium from 2014. Panel A relies on the results of the matched regression (reported in Table 8 column (2)), while the regression was run on the full sample for panel B. Column (1) reports the observed average log wage of non-college educated workers, college educated workers and the difference of the two, the college premium. Column (2) shows the college premium under the counterfactual scenario when all high-novelty innovators conduct only low-novelty innovations. This is calculated by replacing the R&D dummy of all firms to zero, and predicting to wage from the regression for each worker. Finally, column (3) displays the wage premium in the counterfactual scenario when no firms innovate, again predicted from the regression for all workers. Therefore, according to Panel A, the observed college wage premium was 80%, from which 0.3 percentage points was associated with high-novelty innovation and an additional 5.7 percentage points with low-novelty innovation.

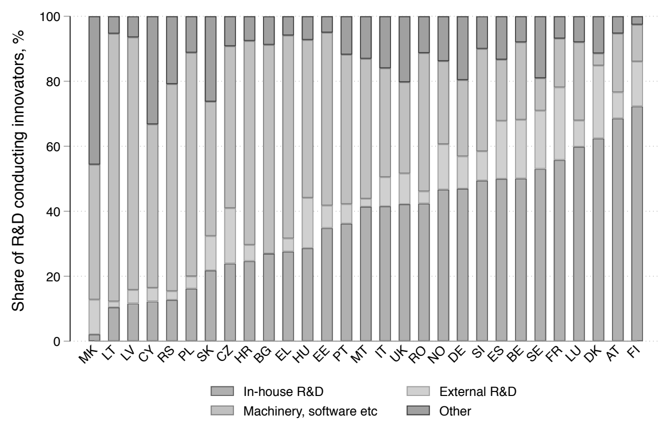
Appendix Figures

Figure A1: The prevalence of low- and high-novelty innovation

Panel A: Share of R&D conducting firms in all innovators



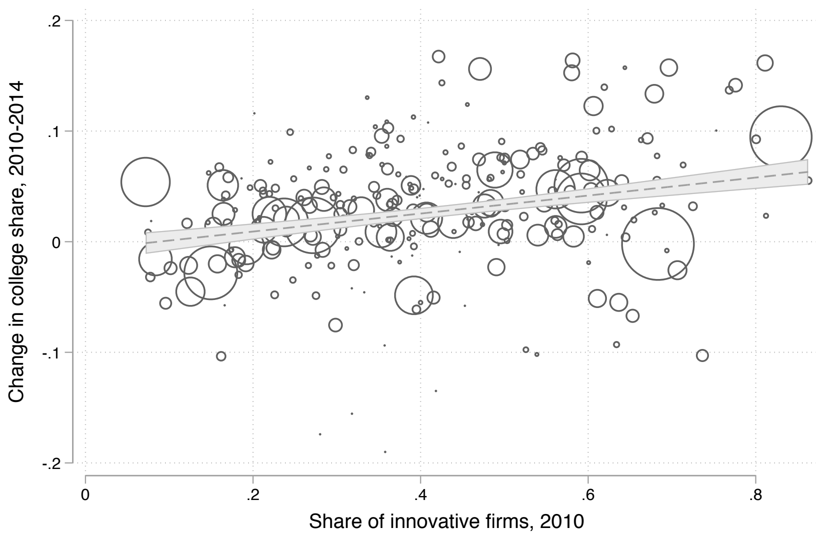
Panel B: Share of R&D in total innovation costs



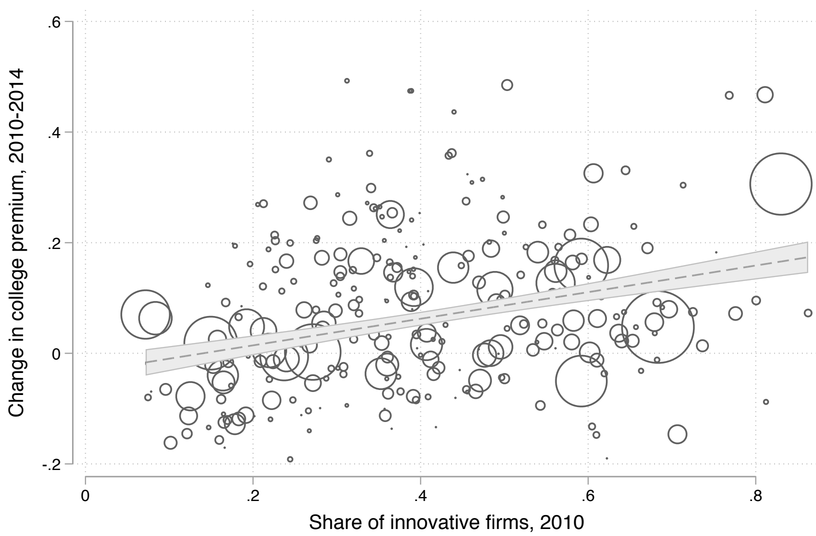
Notes: Panel A of this figure shows the share of firms which conducted R&D relative to all innovative firms by country. Panel B shows the share of different types of innovation expenditures relative to total expenditures on technological innovation. Source: CIS, 2014.

Figure A2: Innovation and subsequent growth in the share and premium of skilled workers at the country-industry level

Panel A: Share innovators and subsequent growth in the share of college-educated employees



Panel B: Share innovators and subsequent growth in the wage premium of college-educated employees



Notes: The figures illustrates the relationship between innovation and subsequent increase in skill demand at the country-1 digit industry level for 25 European countries. In particular, they show how the share of innovative firms (according to the 2010 CIS) is related to the growth in the share of college educated workers (Panel A) and the growth their wage premium (Panel B) between 2010 and 2014, calculated from the Structure of Earnings Survey at the country-industry level. The size of the circles is proportional to the number of firms at that cell in the CIS, and the line shows a weighted regression line with a 95 percent confidence interval.

Appendix Tables

Table A1: Innovation and the college premium: cross-country evidence

LHS: College premium	(1)	(2)	(3)	(4)	(5)
Innovative firms (share)	0.894**	0.907**			0.910**
	(0.408)	(0.411)			(0.430)
R&D firms (share)			-0.329	0.000	-0.046
			(0.315)	(0.348)	(0.320)
Share of college educated		-0.013**		-0.011*	-0.014**
		(0.005)		(0.006)	(0.006)
CEE		0.361***		0.165*	0.364**
		(0.122)		(0.095)	(0.128)
Constant	0.945***	1.432***	1.598***	1.875***	1.463***
	(0.237)	(0.264)	(0.140)	(0.222)	(0.281)
Sample	w/o CEE	all	w/o CEE	all	all
Observations	17	23	16	22	22
R-squared	0.242	0.479	0.072	0.359	0.492

Note: This table reports the cross-country regressions with the college premium as the dependent variable, that underlie the scatterplots in Figure 1. 'Innovative firms' is the share of firms conducting innovation, 'R&D firms' is is the share of R&D conducting firms, CEE is a dummy for new EU member states. Standard errors in parentheses. Significance levels are: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: Linking the datasets: Number of firms in the sample

	(1)	(2)	(3)
Year	CIS	CIS balance sheet	CIS balance sheet structure of earnings
2003	3,950	3,190	1,483
2004	3,950	3,268	1,408
2005	5,094	4,063	2,275
2006	5,094	4,149	1,995
2007	5,390	4,365	1,796
2008	5,390	4,466	2,216
2009	5,120	4,134	1,811
2010	5,120	4,211	1,740
2011	5,482	4,458	1,981
2012	5,482	4,430	2,126
2013	7,243	5,849	2,407
2014	7,243	5,912	2,512
Total	64,558	52,495	23,750

Note: This table shows the number of firms in the sample after the different steps of linking the database. Column (1) shows the number of firms in the CIS in each year. Column (2) shows the number of firms which appear both in the CIS and the balance sheet data. Column (3) presents the number of firms which could also be linked to employee data.

Table A3: Firm innovation status by industry

NACE	Product inn.	Process inn.	Organizational inn.	No innovation	Total
A	3	3	3	8	14
B	21	33	37	294	375
C	3,935	3,373	3,785	7,790	14,436
D	93	239	264	606	988
E	85	257	293	931	1,399
F	49	88	192	643	892
G	267	308	465	1,314	2,183
H	198	337	432	1,203	1,853
I	0	0	0	5	5
J	336	237	371	452	993
L	0	0	3	7	10
M	102	97	139	330	541
N	4	3	13	26	46
Q	...	0	...	...	...
R	0	0	0	...	...
S	0	0	3	9	12
Total	5,093	4,975	5,997	13,611	23,750

Note: This table shows the number of firms conducting different types of innovation in our regressions sample by 1-digit NACE rev 2.2. categories. ...=confidential.

Table A4: Increase in college premium at the country-industry level when controlling for change in skill ratio

	College premium change, 2010-2014			
	(1)	(2)	(3)	(4)
Share of innovative firms (2010)	0.277*	0.220*	0.177	0.240
	(0.143)	(0.112)	(0.130)	(0.147)
R&D-intensity (2010)	-0.019*	-0.003	-0.027**	-0.006
	(0.009)	(0.006)	(0.012)	(0.007)
College share change (2010-2014)	0.060	0.415	0.068	0.555
	(0.605)	(0.499)	(0.495)	(0.460)
country FE		yes		yes
industry FE			yes	yes
Observations	154	152	153	151
R-squared	0.193	0.679	0.304	0.726

Note: These tables shows regressions at the 1-digit industry-country level for 25 European countries. The dependent variable is the change in the share of college educated workers and their skill premium. The main explanatory variable shows the share of innovative firms according to the 2010 CIS wave, measuring innovation activities between 2008 and 2010. This table also controls for the change in college share between 2010 and 2014. Observations are weighted with the number of firms in the country-industry cell from the CIS. Standard errors, clustered at the country level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A5: Innovation and subsequent change in firm-level outcomes: different time periods

	(1)	(2)	(3)	(4)	(5)	(6)
LHS	college wage share	college employment share	college wage share	college employment share	college wage share	college employment share
LHS difference length:	2 years		4 years		6 years	
Innovation	0.008***	0.007**	0.013***	0.015***	0.017***	0.019**
	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.008)
ln capital (d)	-0.006*	-0.010*	-0.008**	-0.012**	-0.006	-0.007
	(0.004)	(0.006)	(0.003)	(0.006)	(0.004)	(0.007)
ln va (d)	-0.007**	-0.014***	-0.010**	-0.009	-0.005	-0.007
	(0.003)	(0.005)	(0.004)	(0.006)	(0.005)	(0.008)
Dependent variable (t-1)	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes
Observations	5,145	5,145	3,357	3,357	2,153	2,153
R-squared	0.060	0.068	0.088	0.098	0.099	0.095

	(1)	(2)	(3)	(4)	(5)	(6)
LHS	TFP (ACF)	TFP (LP)	TFP (ACF)	TFP (LP)	TFP (ACF)	TFP (LP)
LHS difference length:	2 years		4 years		6 years	
Innovation	0.021*	0.027**	0.020	0.028*	0.061**	0.050**
	(0.012)	(0.011)	(0.016)	(0.015)	(0.025)	(0.022)
Dependent variable (t-1)	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes
Observations	5,064	5,064	3,306	3,306	2,122	2,122
R-squared	0.095	0.091	0.129	0.131	0.140	0.148

Note: Note: This table shows the firm-level relationship between innovation and subsequent change of key variables, following Caroli & Van Reenen (2001) and specified in Equation (5). The dependent variables are the long differences of the variables in the column headings, defined as their change between t and $t+2$ in columns (1)-(2), $t+4$ in columns (3)-(4) and $t+6$ in columns (5)-(6). The Innovation dummy shows whether the firm conducted product, process or organizational innovation between years $t-5$ and t, according to the CIS waves conducted in years t and $t-2$. The other two explanatory variables in columns (1)-(3) are long differences (the same length as the dependent variable of log capital stock and log value added. The sample includes firm-years when the subsequent change was available. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A6: Innovation and the college premium: worker-level regressions without skill ratio

LHS: log wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Innovation	0.201***	0.166***	0.165***	-0.026**	-0.024**	-0.005	-0.007
	(0.022)	(0.019)	(0.018)	(0.013)	(0.010)	(0.009)	(0.010)
Innovation x College	0.085***	0.100***	0.085***	0.114***	0.119***	0.083***	0.085***
	(0.027)	(0.023)	(0.022)	(0.013)	(0.013)	(0.020)	(0.021)
Innovation pre-trend					-0.001		-0.007
					(0.020)		(0.014)
Innovation pre-trend x College					0.075***		0.013
					(0.022)		(0.023)
Skill-year FE	yes	yes	yes	yes	yes	yes	yes
Mincer variables		yes	yes	yes	yes	yes	yes
$\hat{p}$			yes	yes	yes	yes	yes
Firm FE				yes	yes	yes	yes
Matched sample						yes	yes
Observations	785,443	785,443	785,443	785,419	785,419	157,714	157,714
R-squared	0.438	0.507	0.517	0.717	0.717	0.699	0.699
Firms	6236	6236	6236	6212	6212	1075	1075
F-value for pre-trend					3.458		12.13
p-value for ptre-trend					0.0630		0.000516

Note: This table re-runs the individual regressions in Table 5 without including the skill ratio. Standard errors, clustered at the firm level are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A7: Innovation and the college premium: innovation variable defined in different time horizons

Innovation dummy from CIS waves:	Current		Current+previous		Current +previous two	
LHS: log wage	(1)	(2)	(3)	(4)	(5)	(6)
Innovation	-0.009	-0.011	-0.009	-0.012	-0.005	-0.007
	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.010)
Innovation x College	0.070***	0.071***	0.072***	0.073***	0.083***	0.085***
	(0.019)	(0.020)	(0.019)	(0.021)	(0.020)	(0.021)
Innovation pre-trend		-0.007		-0.010		-0.007
		(0.013)		(0.014)		(0.014)
Innovation pre-trend x College		0.015		0.014		0.013
		(0.022)		(0.023)		(0.023)
Skill-year FE	yes	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes	yes
$\hat{p}$	yes	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes	yes	yes
Observations	142,249	142,249	154,024	154,024	157,714	157,714
R-squared	0.697	0.697	0.699	0.699	0.699	0.699

Note: This table shows that the main results from the individual regressions are not sensitive to the time period used to define the innovation dummy. In columns (1) and (2), the firm is only considered innovative if it conducted innovation according to the current CIS wave. In columns (3) and (4), the firm is considered innovative if it conducted innovation either in the current or the previous wave. Columns (5) and (6), our preferred version, define the innovation variable based on the current and the two previous waves. All regressions represent the preferred specification, Table 5 column (6). Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A8: Innovation and the college premium: incumbent workers

LHS: log wage	(1)	(2)	(3)	(4)	(5)
Innovation	0.202***	0.176***	0.174***	-0.020**	-0.010
	(0.023)	(0.020)	(0.020)	(0.009)	(0.010)
Innovation x College	0.087***	0.097***	0.082***	0.101***	0.076***
	(0.028)	(0.025)	(0.023)	(0.013)	(0.020)
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables		yes	yes	yes	yes
$\hat{p}$			yes	yes	yes
Firm FE				yes	yes
Matched sample					yes
Observations	576,496	576,496	576,496	576,446	114,644
R-squared	0.449	0.497	0.509	0.724	0.705

Note: This table shows that incumbent workers' wages change very similarly to the wages of all workers. This table repeats the individual-level regressions in Table 5 for the subset of workers who had been working at the firm for more than 24 months. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A9: Innovation and the college premium: switcher workers and worker fixed effects

LHS: log wage	(1)	(2)	(3)	(4)
Innovation	0.031**	0.001	0.002	0.003
	(0.013)	(0.008)	(0.008)	(0.008)
Innovation x College	0.044**	0.044***	0.030*	0.023**
	(0.021)	(0.017)	(0.017)	(0.010)
Skill-year FE	yes	yes	yes	yes
Firm FE		yes	yes	yes
incumbent sample			yes	
$\hat{p}$			yes	yes
worker FE				yes
Observations	36,011	36,011	31,289	36,011
R-squared	0.426	0.728	0.740	0.948
Clusters	699	699	679	699

Note: This table re-runs regressions from Table 5 on a sample of workers which are observed in the firm both before and after an innovation. Column (4) also includes workers fixed effects. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A10: Innovation and different worker-level outcomes

	(1)	(2)	(3)	(4)
LHS:	total wage	base salary	got bonus	log hours
innovation	-0.005	-0.019	0.020	-0.000
	(0.009)	(0.012)	(0.014)	(0.002)
innovation x college	0.083***	0.095***	-0.008	0.000
	(0.019)	(0.025)	(0.017)	(0.003)
Skill-year FE	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes
$\hat{p}$	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes
Observations	157,714	105,081	157,714	157,714
R-squared	0.697	0.694	0.442	0.700

Note: This table shows the results of worker-level regressions, described in Equation (6), with different dependent variables. All the regressions follow the preferred specification from Table 5, column (6). For a reference, column (1) repeats the regression with total wage as the dependent variable. In column (2), the dependent variable is the base wage without bonuses and other flexible wage elements. Column (3) estimates how innovation is related to the probability of receiving any bonus. Finally, column (4) estimates whether innovation leads to a chnage in hours worked. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A11: Innovation and the college premium: routine jobs

RHS: log wage	(1)	(2)	(3)	(4)	(6)
Innovation	0.215***	0.179***	0.177***	-0.021**	-0.006
	(0.018)	(0.016)	(0.016)	(0.011)	(0.010)
1-RTI	0.055***	0.033***	0.030***	0.049***	0.059***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.007)
Innovation x college	0.040	0.059***	0.046**	0.091***	0.079***
	(0.025)	(0.022)	(0.021)	(0.013)	(0.020)
Innovation x (1-RTI)	0.050***	0.047***	0.046***	0.025***	0.002
	(0.011)	(0.008)	(0.008)	(0.006)	(0.008)
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables		yes	yes	yes	yes
$\hat{p}$			yes	yes	yes
Firm FE				yes	yes
Matched sample					yes
Observations	784,732	784,732	784,732	784,708	157,638
R-squared	0.456	0.517	0.526	0.725	0.705
Clusters	6236	6236	6236	6212	1075

Note: This table re-runs the individual regressions in Table 5 but also includes a proxy for (non) routine skills used in the job, following Autor et al. (2003) and its interaction with the college dummy. Note that the variable in the regression is increasing in the share of non-routine tasks. Standard errors, clustered at the firm level are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A12: Innovation and the skill premium: polarization?

LHS: log wage	(1)	(2)	(3)	(4)	(5)
Innovation	0.194***	0.167***	0.165***	-0.027*	-0.018
	(0.035)	(0.031)	(0.030)	(0.015)	(0.011)
Innovation x Primary	-0.015	-0.031	-0.030	-0.009	0.023
	(0.035)	(0.030)	(0.030)	(0.011)	(0.017)
Innovation x Vocational	0.022	0.018	0.017	0.006	0.021*
	(0.027)	(0.025)	(0.025)	(0.008)	(0.013)
Innovation x College	0.092***	0.101***	0.086***	0.114***	0.095***
	(0.032)	(0.026)	(0.026)	(0.013)	(0.021)
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables		yes	yes	yes	yes
$\hat{p}$			yes	yes	yes
Firm FE				yes	yes
Matched sample					yes
Observations	785,443	785,443	785,443	785,419	157,714
R-squared	0.438	0.503	0.513	0.715	0.697

Note: This table investigates whether innovation is associated with the polarization of wages by distinguishing between four education categories rather than only non-college/college. To this end, the main individual-level wage regressions in Table 5 are augmented with the innovation dummy interacted with 3 education levels, leaving secondary education as the base category. The interactions show innovative firms' premia for each education category relative to the premium of workers with a secondary degree. Polarization could mean that that both innovation x primary and innovation x college variables are significant, i.e. the innovation premia of low and high-skilled workers is larger than that of workers with a medium level of education. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.