Technological Change and Skill Demand in Non-Competitive Labor Markets *

Attila Lindner
University College London

Balázs Muraközy
CERS, Hungarian Academy of Sciences

Balázs Reizer
CERS, Hungarian Academy of Sciences

Ragnhild Schreiner
University of Oslo

March 2020

Abstract

This paper investigates the consequences of technological change in the presence of non-competitive labor markets. We propose a model of technological progress where firms invest in innovation in the hope of developing new technologies. A successful innovation elevates firm-level labor demand, and so firms have to raise wages to hire more workers. Unlike in models where wages are set competitively, in this framework firm-level wage responses reveal information about the nature of technological change. We show that one can infer the extent which technological change is skill biased by jointly investigating the effect of innovation on the firm-level skill ratio and on the skill wage premium. We apply this idea by exploiting unique firm-level innovation surveys linked to employee-employer data from Hungary and Norway. We show that firm-level technological change raises the skill ratio and also the skill premium in both countries. The increase in the skill-premium is not driven by the change in composition of the workforce and, in line with the predictions of the non-competitive labor markets, wages of new entrants are also affected. Both high- (e.g. R&D based) and low-novelty value innovations are equally skill biased. Among low-novelty innovation types, technological innovation are the most skill-biased, while organizational innovation is less so.

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. This research has been funded by the European Union's Horizon 2020 research and innovation programme, grant agreement No 822390 (MICROPROD). 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 A2). 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.

To examine the relationship between technology and firm-level skill demand we develop a model of technological progress where firms invest resources to explore new ideas in the hope of developing

[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.

or adopting technologies that are new to the firm (but not necessarily to the market). A successful innovation leads to a change in production function. The essence of this model is similar to Aghion & Howitt (1992). Nevertheless, we deviate from that model by introducing non-competitive wage setting along the lines of Card et al. (2018) and Lamadon et al. (2018). Wages in our framework are not any more exogenously given, but expanding firms need to set higher wages (Card et al. 2018).

We model skill demand in a framework where firms have two inputs in the production function: high skilled labor and low skilled labor. The basic insight from the model is that if wages are set in a non-competitive environment, then a negative relationship emerges between relative skill demand and relative wages at the firm level as long as there is no skill-biased technological change. Therefore, when a Hicks-neutral technological change is introduced, either relative demand falls and relative wages increase or vice versa. In contrast, an increase in both relative demand and relative wages in response to technological change can only be consistent with a skill biased change. The intuition behind this result is exactly the same as the one in Katz & Murphy (1992) who examined the relationship between relative supply and relative demand. We highlight that when labor markets are non-competitive, this relationship emerges at the firm-level as well.

Our framework also highlights that, unlike in models with competitive wage settings (see e.g. Caroli & Van Reenen 2001, Bøler 2015), in the presence of non-competitive wage setting even a Hicks-neutral shock can affect the skill ratio when firms have 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 their increased demand drives up wages to a smaller extent. This implies that focusing only on the skill ratio 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.

We test the key insights of the model by using exceptionally rich microdata from Hungary and from Norway. These two countries are at a very different distance from the technological frontier. Most Hungarian firms adopt technologies developed in Western Europe. At the same time, Norway itself develops new technologies and those are adapted later by poorer countries, which are further from the frontier. There labor markets are also markedly different. As a result, understanding the effect of the technological change in countries that are so different from each other is likely to provide a more complete picture on the impact of technological change on wage inequality.

In both countries we rely on the rich information 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). The survey captures all significant changes in the firm's production function and so the innovation concept in the data has a clear counterpart in the theoretical model.

We estimate how the skill premium is related to innovation by implementing 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 make sure that our estimates are not driven by compositional change in the quality of workforce after innovation, we include firm and worker fixed effects when estimating on the Norwegian data. Unfortunately, such a strategy is not feasible in Hungary because of data limitations. There we match firms based on observable characteristics in the initial period. In both cases we can restrict the attention to workers who were at the firm before innovation started, and get an estimate on them to make sure that the wage premium is not driven by the changes in composition. We find that starting to innovate is associated with a 5-8 percentage point increase in the wage premium in Hungary and 4-6 percentage point increase in Norway.

These estimates are robust to controlling for industry-skill-year fixed effects, occupation-year fixed effects, not sensitive to alternative timing assumptions, and are not driven by pre-innovation wage differences. 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 in addition to the increasing college premium.

Note that a substantial change in the wage premium is an important piece of evidence in itself that labor markets are non-competitive. Nevertheless, if there is imperfect substitutability between incumbent and new hires then incumbent workers can extract rents from the firm in the form of wage premia even if labor markets are perfectly competitive (Becker 1964, Kline et al. 2018). In contrast, when the increase in wage premium is a result of upward-sloping firm-level labor supply, one can expect that the firm will also have to pay more for new entrants. In line with this latter prediction we find that both new entrant and incumbent workers experience an increase in their wage premium following an innovation.

To assess the change in skill ratio we follow 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 is associated with subsequent growth both in the employment and wage share of college-educated workers and subsequent productivity growth. The joint increase in skill-ratio and skill wage premium suggests that technological changes are skill biased in Hungary and in Norway.

After establishing that innovation is skill biased both in Hungary and Norway, we study the heterogeneity of innovation along two dimensions.[4] 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 than high novelty innovation.

[4] These results are only produced for Hungary so far.

Second, we distinguish between technological and organizational innovation. Similarly to (Caroli & Van Reenen 2001) we find that all these types of innovation appear to increase the skill ratio to some extent. Nevertheless, we find that the skill premium is mainly driven by technological innovation, while organizational innovation has only a minor, non-significant impact on skill premium. These evidence suggest that technological adoption is skill biased while contrary to conclusion of (Caroli & Van Reenen 2001), the lack of evidence on skill premium suggest that organizational innovation is unlikely to be skill biased.

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. The richness of the data also allows us to provide direct evidence on worker-level wages and control for compositional change by including worker fixed effects in the regression.

Furthermore, our model of technological progress also highlights 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. In fact, in some important cases (e.g. organizational change), we find quantity response, but not wage responses. Those examples highlight that making conclusions solely based on quantity responses can be misleading.

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 the theoretical 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

2.1 Setup

We present a model of technological progress through firm-level innovation activities. We apply the basic idea of Aghion & Howitt (1992) and assume that firms need to invest resources to upgrade their production function. The probability of innovation will depend on various factors: the size of this investment, the distance from the frontier technology, the future expectations about output demand and the evolution of prices, and some luck. Our key innovation is that we introduce a non-competitive wage setting process following Card et al. (2018): firms that want increase their employment after a successful innovation need to set higher wages. In our model therefore both wages and employment are set endogenously.

Let us start with describe the firm's and worker's problem and examine how firm-level technological change affect employment and wages. Later we will endogenize technological change and describe that under what assumptions we can identify whether it is skill-biased.

There are continuum of firms operating in a local labor market. Firms use two inputs for production: high skilled labor (H_{jt}) and low skilled labor (L_{jt}). In every period firms produce with the following CES technology:

$$y_{jt} = A_{jt} \left[\theta_{jt} H_{jt}^{\rho} + (1-\theta_{jt}) L_{jt}^{\rho}\right]^{\frac{1}{\rho}} \tag{1}$$

The production function is represented by three technology parameters: A_{jt} is the Hicks-neutral revenue productivity term, θ_{jt} represents the extent which the technology used by the firm is skill biased, and $\rho \leq 1$ govern the elasticity of substitution between skilled and unskilled labor. The elasticity of substitution between skilled and unskilled labor is $\sigma = 1/(1-\rho)$. Importantly, successful innovation affect the technology parameters.

Firms maximize their profit given this production function:

$$\pi_{jt}(A_{jt}, \theta_{jt}) = \max_{w_{L_j}, w_{H_j}} p_{jt}y_{jt} - H_{jt}(w_{H_{jt}})w_{H_{jt}} - L_{jt}(w_{L_{jt}})w_{H_{jt}} \tag{2}$$

For the sake of brevity, we assume that the firms are price takers at the product market. This does not affect any of the results presented here, but since we want focus on price setting process on the input side, we abstract away from price setting on the output market.

We follow recent developments in modeling labor supply and use a discrete choice framework to model workers' choice of firms (Card et al. 2018, Lamadon et al. 2018). We assume that each firm posts a pair (w_{Ljt},w_{Hjt}) of skill-specific wages that all workers costlessly observe. For workers 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 under these assumptions the approximate firm-specific upward-sloping labor supply functions are:

$$\begin{aligned} L_j(w_{Lj}) &= \ln(\mathcal{L}_{L_t}) + \beta_L \ln w_{Lj} + a_{Lj} \\ \ln H_j(w_{Hj}) &= \ln(\mathcal{H}_{H_t}) + \beta_H \ln w_{Hj} + a_{Hj} \end{aligned} \tag{3}$$

where $\ln(\mathcal{L}_{L_t})$ and $\ln(\mathcal{H}_{H_t})$ are time-varying constants common to all firms in the market. Note that as β_L, $\beta_H \to \infty$, these supply functions become perfectly elastic and we approach a competitive labor market. Firms maximize their profits by taking into account that they face upward sloping labor supply functions.

Technology is determined endogenously in the model. Firms need to invest resources to develop better technologies. More specifically, if the firm spends z_{jt} dollars on innovation inputs, then it will successfully innovate with probability $\lambda\left(\frac{A_{jt}}{\widetilde{A_t}}\right)z_{jt}$, where $\widetilde{A_t}$ is the frontier productivity level at time t. We want to emphasize that innovation spending, z_{jt}, does not solely represent R&D activities. In fact, firm level technological change often simply means adopting an existing technology, and so in our model even firms far away from the frontier will innovate. Note that the probability of discovering, can depend on how far the firm is from the technological frontier. A successful innovation improves the TFP of the firm, from A_{jt} to γA_{jt}.[5] At the same time we allow innovation to affect increase θ_{jt}, the extent which the new technology is skilled biased to $\kappa\theta_{jt}$.

The firms' problem is the following:

[5]Given the broad range of innovation activities we consider, we remain agnostic on the exact way innovation affects A_{jt}, which is a revenue TFP empirically. In particular, some innovations may increase physical productivity, while others improve the quality of the product, and, therefore, enable the firm to charge a higher price.

$$V_{jt}(A_{jt}, \theta_{jt}) = \max_{z_{jt}} \pi_{jt}(A_{jt}, \theta_{jt}) - z_{jt} + + \left[\lambda \left(\frac{A_{jt}}{\widetilde{A_t}} \right) z_{jt} V_{jt}(\gamma A_{jt}, \kappa \theta_{jt}) + \left(1 - \lambda \left(\frac{A_{jt}}{\widetilde{A_t}} \right) z_{jt} \right) V_{jt}(A_{jt}, \theta_{jt}) \right] \tag{4}$$

where R is the market-level discount factor.

2.2 Model

The model predicts that firms invest into innovation inputs up the point where the marginal benefit of innovation investment is the same its marginal cost (w_z), formally:

$$w_Z = R \left[\lambda \left(\frac{A_{jt}}{\widetilde{A_t}} \right) \left(V_{jt}(\gamma A_{jt}, \kappa \theta_{jt}) - V_{jt}(A_{jt}, \theta_{jt}) \right) \right]$$

If the expected benefit of innovative investments, formally $V_{jt}(\gamma A_{jt}, \kappa \theta_{jt}) - V_{jt}(A_{jt}, \theta_{jt})$ is larger, then firms put more resources into innovation and so they are more likely to innovate. The benefits of innovation can depend on the local labor market characteristics and the extent which innovation is skilled biased.

Now let us solve the firms' profit maximization problem. It is easy to show that the profit maximization problem leads to the following expression on wage ratio:

$$\ln \frac{w_{H_{jt}}}{w_{L_{jt}}} = \ln \frac{\beta_H}{1+\beta_H} - \ln \frac{\beta_L}{1+\beta_L} + \ln \frac{\theta_{jt}}{1-\theta_{jt}} - \frac{1}{\sigma} \ln \frac{H_{jt}}{L_{jt}} \tag{5}$$

This equation resembles the key equation in the skill biased technological change literature that shows the relationship between relative demand and relative wages of college and non-college workers (see e.g. Violante 2008, Katz & Murphy 1992). Nevertheless, here the relationship emerge at the firm-level, since firms have some wage setting power.

Equation (1) also highlights that the relationship between relative wages and relative demand must be negative if the technological change is not skilled biased (θ_{jt}). Consequently, showing that relative demand and relative wages are both positively affected by technological change implies that the change was skill biased. This point is made in the seminal paper by Katz & Murphy (1992) who examined aggregate level supply and demand factors. Nevertheless, here we show that a similar exercise can be done at the firm-level when firms operating in non-competitive labor markets.

It is also worth expressing the relative demand in the model:

$$\ln \frac{H_{jt}(w_{H_{jt}})}{L_{jt}(w_{L_{jt}})} = \frac{\sigma \beta_H}{\sigma + \beta_H} \left(\ln A_{jt} + \frac{1}{\sigma} \ln y_{jt} + \ln \theta_{jt} + \frac{1}{\beta_H} a_{H_{jt}} + \Lambda_H \right) + - \frac{\sigma \beta_L}{\sigma + \beta_L} \left(\ln A_{jt} + \frac{1}{\sigma} \ln y_{jt} + \ln(1-\theta_{jt}) + \frac{1}{\beta_L} a_{L_j} \right. \tag{6}$$

where $\Lambda_H = \ln \frac{\beta_H}{\beta_H+1} + \frac{1}{\beta_H} \ln\left(\mathcal{H}_{H_t}\right)$ and $\Lambda_L = \ln \frac{\beta_L}{\beta_L+1} + \frac{1}{\beta_L} \ln\left(\mathcal{H}_{L_{jt}}\right)$ are constants. It is easy to show that the effect of Hicks-neutral technological change on relative skill demand is given by the following equation:

$$d \ln \frac{H_{jt}(w_{H_{jt}})}{L_{jt}(w_{L_{jt}})} = \left(\frac{\sigma\beta_H}{\sigma+\beta_H} - \frac{\sigma\beta_L}{\sigma+\beta_L} \right) \left(d\ln A_{jt} + \frac{1}{\sigma} d\ln y_{jt} \right) \tag{7}$$

where $d\ln y_{jt} > 0$. This equation highlights that the effect of Hicks-neutral technological change on firm-level skill ratio can be positive or negative depending on the relative size of β_H and β_L. To understand the intuition of this result suppose that $\beta_H > \beta_L$ and so the high skilled labor market is more competitive, with firms having less wage setting power in that market. A firm that experiences a positive Hicks-neutral shock will want produce more, which requires more workers. The type of workers it hires depends on the relative elasticity of supply on the two markets. If the high skilled market is more competitive, it is cheaper for the firms hire high skilled workers as it can be done with smaller wage increase and so firms will shifts its demand toward those type of workers. The intuition is similar to the problem of a firm which a firm faces when it sells to markets with different elasticities of demand: following a fall in its costs, it will increase its sales more in the more competitive market.

This discussion also highlights a major limitation of the existing literature that solely focus on the relationship between firm-level skill ratio or the spending ratio to assess skill bias (see e.g. Caroli & Van Reenen 2001). As equation (7) shows, examining alone the change in skill ratio is not sufficient when labor markets are non-competitive . Instead, similarly to the macro literature, we propose here to jointly look at the relationship between relative demand and relative wages.

Finally, in our model technological change is an outcome of endogenous firm-level decisions. It is possible that innovative investment is driven by expectations about future prices, p_{jt}, or about future labor market characteristics. For instance, firms might anticipate that low skilled workers will be less abundant and so they are more likely to innovate. Nevertheless, it is worth emphasizing that this source of endogeneity in the innovation decision does not fundamentally alter the key conclusions of Equation (5): in absence of any change in θ_{jt} the firm-level relationship between relative wages and relative demand will be negative. In this particular example, whenever low skilled workers are less abundant and so relative demand increases, we expect to see a relative fall in wages in absence of any change in θ_{jt}. As a result, evidence that technological change jointly increase the relative demand and the relative wages can only be reconciled with a shift in θ_{jt}.

2.3 Applying the model

Our empirical investigation will examine the consequences of introducing a successful innovation, which, according to the model, can be characterized by the Hicks-neutral and the skill biased technological change it triggers, or the (γ, κ) pair.

The main question of the empirical investigation is whether the innovation is skill biased, e.g. $\kappa > 1$. According to our discussion, the appropriate way to test this hypothesis is to jointly investigate the innovation's effect on the firm-level skill and wage premium. If we find that innovation raises both of these quantities, then the innovation must be skill biased.

Note that the investigation of the firm level skill premium also test indirectly whether there is a need to consider non-competitive labor markets when modeling the consequences of technological change. In fact, a substantial change in the wage premium is an important piece of evidence in itself that labor markets are non-competitive. Wage effects, however, can be present in competitive labor markets when firms share rents with workers. However, in the rent-sharing literature the rent is typically distributed among incumbent workers (Kline et al. 2018). In contrast, when the increase in wage premium is a result of upward-sloping labor supply, one can expect that the firm will also have to pay more for new entrants. Empirically, we can investigate this by estimating the effect of the innovation separately for new entrant and incumbent workers. A significant effect on new entrants is in line with our framework.

*******TODO

Identification assumption

Use this equation to show how different types of endogeneity affect the estimates

Suppose that the firm knows the parameters of the innovation, θ and κ and decides on whether to introduce one under different types of expected shocks.

Shock (i): profitability shock. Now p_{jt} in equation 2 increases. This will indeed cause endogeneity when estimating the TFP effect of innovation, γ, but will not be a problem for estimating this skilled biased effect, κ

Shock (ii): general labor supply shock. This is a proportional increase (defined appropriately) in both $\ln(\mathcal{L}_{L_t})$ and $\ln(\mathcal{H}_{H_t})$. Again, it will not affect the estimation of κ

Shock (iii): type specific labor supply shock: only $\ln(\mathcal{L}_{L_t})$ (or $\ln(\mathcal{H}_{H_t})$) changes. This may indeed be a problem for estimating κ, but will work against us

Conclude explicitly that explicitly that level change biases TFp effect but not the relative price effect

****

Also, different types of technological change can have different (γ, κ) parameters. For example, technological innovation may have different effects than organizational innovation. Or, high novelty innovation may involve a different skill biased part than lower novelty innovation. Innovation surveys provide ample information on the type of innovation, therefore we can test whether only some of these are skill biased. Similarly, innovation may affect the production function differently if the country is closer to the technology frontier. We can test such differences by comparing the effects of innovation in Hungary and Norway.

3 Data

Our work is based on the Community Innovation Survey for Hungary and Norway, as well as the Structure of Earnings Survey and Balance Sheet data for Hungary, and the Employer-employee register merged with Balance Sheet data for Norway.

3.1 Innovation data

The first data source is the Hungarian and Norwegian versions 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 has used the CIS 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 firms' innovation activities during the previous three. In this paper we use six waves of the CIS survey from the period between 2004 and 2014 (five waves: 2004-2012 for Norway). In both countries, the sample size has been progressively increasing from about 4,000 firms in 2004 to more than 7,000 in 2014 (Table A3).

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 which can be classified into 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.[6]

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 one 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).[7] We create a dummy variable indicating whether the firm reports a product and process innovation developed by the firm itself. 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 three 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 **??**.

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

[7]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 Matched employer-employee data

We link the CIS data to employer-employee data from both countries. The source is the Structure of Earnings survey in Hungary, which is a repeated cross sectional survey. In Norway, we rely on the employee register, which is a panel dataset of all employees.

3.2.1 Hungary: Structure of Earnings Survey and administrative Balance sheet data

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.[8]

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.

For important firm-level variables, including the number of employees, industry classification, ownership and key financial variables, we complement these data with administrative balance sheet 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.

[8] 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.

The three Hungarian databases can be linked based on unique firm identifiers.[9]. 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.[10]

Table A3 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 A4 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%).

3.2.2 Norway: Employer-employee register

The employer-employee register, provided by Statistics Norway, contains all employment spells in Norway. This data set includes information on wages and days worked per employment spell per year. We merge the employer-employee register to data on worker demographics that includes information on level of education, age and gender. Finally, these data are matched to firm balance sheet data for all limited liability companies in Norway, where we extract information on value added and capital.

For the worker level analyses, we start out with the employer-employee register for the years 2002-2013 and keep the main (highest annually paid) employment spell of full-time workers aged 16 to 67. To be included in the data we further require that the worker is employed in a firm for at least 30 days a given year. Finally we drop the (few) workers for whom demographic information on age, gender and level of education is missing. This gives an unbalanced panel data set containing information on 1,004,812 workers employed in 118,512 different firms over the 12-year period 2002-2013. This data set is merged to five waves of the CIS survey for Norway that covers an unbalanced sample of 15,857 firms over the years 2004-2012. Only the workers employed in firms that at some point in time are included in the CIS will contribute to identifying the effect of innovation on the

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

[10] 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.

college wage premium. However, we make use of the full universe of workers and firms to estimate worker and firm fixed effects.

For the firm level analyses, data on firm average daily wage rates for college and non-college educated workers, as well as the number of employees in a firm, come from aggregating up from the worker level sample.[11] However, in the firm level analysis, we keep only firms that are sampled in the CIS, and at the same time are limited liability firms for whom we have data on value added and capital.

4 Empirical approach

Based on our morel, the main aim of the empirical analysis is to estimate the effect of innovation on the skill ratio and the skill premium to test whether innovation is skill-biased. We use a firm-level diff-in-diff strategy to estimate the effect of innovation on the skill ratio, following Caroli & Van Reenen (2001). Worker-level regressions yield themselves better to estimating the effect of innovation on the skill premium, because such regressions enable us to control for worker characteristics. The worker level regressions also allow us to test whether innovation affects only incumbents (as some rent sharing models predict) or also new entrants, as predicted by the wage setting framework. A further question is which kind of innovation is skill biased, which we investigate by including different innovation variables into these regressions. Finally, we use a simple decomposition to quantify the importance of low- and high-novelty innovation in the aggregate cross-sectional skill premium.

4.1 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:

$$\Delta y_{jt} = \beta * innovation_{jt} + \gamma * \Delta X_{jt} + \delta * y_{jt-1} + \eta_{st} + \epsilon_{jt} \tag{8}$$

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 $t+6$.[12] $innovation_{jt}$ 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.[13] 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.[14]

[11] In the firm level analysis part-time workers are included.
[12] We Winsorize this and the other long difference variables at the 5th and 95th percentiles.
[13] We estimate the TFP change by the ACF procedure, which already takes into account the change in inputs.
[14] 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.

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.

4.2 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:

$$\ln wage_{ijt} = \beta_u \times innovation_{jt} + \beta_s \times innovation_{jt} \times college_{it} + \gamma \times X_{ijt} + \\ + \eta_i + \varphi_j + \varsigma_{st} + \varepsilon_{ijt} \quad (9)$$

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 fixed effects, eta_i are worker fixed effects while ς_{st} are skill-year fixed effects, for the 4 categories of schooling.[15] 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. There are a number of challenges to giving β_s in Equation (9) a causal interpretation. One issue is firm heterogeneity in terms of 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. However, 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. Therefore, we check whether such pre-trends are present by including dummies indicating whether the firm will innovate in the next wave.

Another important issue is the self-selection of workers based on unobservables. Higher productivity workers may be more likely to select to innovative firms. To make sure the results are not driven by changing worker composition correlated with innovation, we include worker fixed effects

[15] 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.

when estimating on the Norwegian data.

Unfortunately, such a strategy is not feasible in Hungary, where the Structure of Earnings survey does not have a worker identifier to link observations across waves. Instead, we do a matching at the firm level to create a control group to innovative firms which is as similar as possible. 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 them from firms which do not.

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.[16] 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[17] 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.

A further 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.

[16] 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.

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

4.3 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 (9):

$$\begin{aligned}\ln wage_{ijt} = \beta_u \times innovation_{jt} + \beta_s \times innovation_{jt} \times college_{it} + \\ + \beta_u^n \times innovation_{jt} \times novelty_{jt} + \beta_s^n \times innovation_{jt} \times college_{it} \times novelty_{jt} + \\ + \eta_i + \gamma \times X_{ijt} + \varphi_j + \varsigma_{st} + \varepsilon_{ijt} \quad (10)\end{aligned}$$

Our main parameter of interest is still β_s, which shows the effect of low-novelty innovation on the college premium in this setting.[18] 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 (9).[19]

4.4 Decomposition of the aggregate skill premium

The results of regression (10) 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.

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 (10) 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

[18] 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.

[19] 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.

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 (10).

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

In this section, we first describe both firm-level results, with a focus on the skill share and the at the worker level, with a focus on the skill premium. In Appendix 2 we present further evidence at the more aggregated country-industry level.

5.1 Firm-level evidence

Table 1 compares innovative and non-innovative firms in the two countries. Two types of differences are apparent. First, in line with much of the literature (Griffith et al. 2006), innovative firms are larger and more capital intensive in both countries and also more productive in Hungary. Second, innovation is indeed associated with higher skill levels. In particular, both the average years of education and the share of college graduates are substantially higher.

Table 2 shows the role of between-firm variation in total wage variation in Hungary, following the decomposition of Song et al. (2015). Firm fixed effects explain about 50% of wage variation, which is relatively large in international comparison.[20] 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 main firm-level regression results are presented in Table 3 for the two countries. We start by presenting results for the impact of innovation on 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 of skilled workers increase by 1.7 percentage points in Hungary and 0.9 percentage points in Norway during the six-year period following firm innovation. Column (2) shows that this increase is, almost entirely, explained by the increasing employment share of these workers. Column (3) shows that innovation is associated with stronger employment growth,

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

with a significant estimate in Norway. Finally, Column (4) confirms that innovation is associated with stronger subsequent productivity growth, in the order of 1 percentage point per year.[21]

The main takeaway from these results is that, in fact, innovation raises the share of high-skilled workers. The results are remarkably similar for Hungary and Norway in this respect.

5.2 Worker-level evidence

The main individual-level results for Hungary and Norway are presented in Panel A and B of Table 4. Column (1) shows results on the full sample when only skill-year fixed effects are included. According to these results, in Hungary, workers without a college degree earn 20.1 percent more in innovative firms (relative to workers with similar education levels in non-innovative firms), while this difference is 28.6 percent for college educated workers (compared to college educated workers in non-innovative firms). The cross sectional wage premium of innovative firms is somewhat smaller in Norway, with low and high skilled workers earning 10.7 and 16.4 percent more in innovative firms, respectively. In column (2) we also control for worker observables. In general, Mincer variables do not explain much of the innovative firm wage premium.

We also include firm fixed effects in column (3). In this specification, the low-skilled innovation premium becomes negative in both countries, while the college innovation premium becomes even higher than before, at 10.5 and 12.3% 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.

Our preferred specifications are reported in column (4). In Norway, the structure of the data allows us to include worker fixed effects, while in Hungary, we do matching at the firm level. Importantly, while the estimates become smaller in both countries, they remain highly significant both in economic and statistical terms. In Hungary, high-skilled employees experience a 6.7% increase in their salary following a successful innovation, while this effect is 4% in Morway. The fact that matching at the firm level (using the Hungarian data) changes the estimates in a similar way to including worker fixed effects (using the Norwegian data) suggests that the results presented in Column (4) are unlikely to be driven by changes in worker composition.

To sum up, we find remarkably similar results in the two countries, with a 4-6% increase in college educated workers' wage premium following an innovation. Together with the firm-level results that, for both countries, demonstrate an increase in the skilled share following innovation, we conclude that innovation seems to be skill biased both in Hungary and Norway.

A key characteristic of our framework is that the increase in the wage premium reflects that the firm has to pay higher wages when hiring new workers following the innovation. This implies that, in

[21] A potential concern with the long difference specifications is that their results may be sensitive to the choice of the length of the difference. When re-running the main regressions with 2, 4 and 6-year long lags in Hungary we find a positive association between innovation and subsequent growth both in college share and TFP. The effects are stronger for longer differences suggesting that innovation has prolonged effects.

contract to some rent sharing models (e.g. Kline et al. 2018), new hires, rather than only incumbent workers, should also receive a higher skill premium. Table 5 reports results for incumbents (working at the firms for at least 24 months) and new entrants (other workers) separately. In particular, the regressions include interactions of the innovation and the college variables (as well as their interaction) with the new entrant and incumbent dummies. The results show that both new entrant and incumbent high skilled workers receive a higher premium following an innovation in both countries. The results are robust to all the different specifications. This finding is strongly in line with the assumption that wage setting power is key in explaining the effect of innovation on the wage premium.

A number of robustness checks are presented in Table 6, all starting from our preferred specification (column (4) of Table 4. First, column (1) includes (1-digit) 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.[22] Another potential concern is that the construction of our innovation variable, based on the current and two previous waves, does not capture the timing of the effects of innovation on the wage premium adequately. We investigate the importance of this issue in columns (3) and (4). In column (3) we define the innovation variable based only on the corresponding CIS wave and in column (4) based on the corresponding and the previous wave. Finally, in column (5) we include an innovation pre-trend dummy variable indicating that the firm will innovate in the following wave, and its interaction with the college dummy. All in all, the main results are robust to all these alternative specifications.

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 7 we include (in the case of Hungary) the measure proposed by Autor et al. (2003) and its interaction with innovation besides college and its interaction.[23] 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 A6 reports results (again in the case of Hungary) 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

[22]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.

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

categories do not seem to change after innovation takes place, while the wages of college educated workers do increase substantially.

5.3 Heterogeneity of innovation

Let us start investigating heterogeneity with firm-level regressions, reported in Table 8. 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 (10). 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 9 estimates Equation (10) 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 10 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.

5.4 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 **??** 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 A5 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.

5.5 Quantifying the importance of innovators in the aggregate skill premium

Panel A of Table 11 presents the results of the simple decomposition exercise described in Section 4.4 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 9. 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 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 (10) 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 rich dataset which combines detailed firm-level information on innovation with worker-level wage data from Hungary and Norway. 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 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 play less important role.

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'.

Aghion, P. & Howitt, P. (1992), 'A model of growth through creative destruction', Econometrica **60**(2), 323–51.
URL: *https://EconPapers.repec.org/RePEc:ecm:emetrp:v:60:y:1992:i:2:p:323-51*

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.

Becker, G. (1964), 'S. human capital: A theoretical and empirical analysis, with special reference to education'.

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.

Katz, L. F. & Murphy, K. M. (1992), 'Changes in relative wages, 1963–1987: supply and demand factors', The quarterly journal of economics **107**(1), 35–78.

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.

Lamadon, T., Mogstad, M. & Setzler, B. (2018), Earnings dynamics, mobility costs, and transmission of market-level shocks, in 'Society for Economic Dynamics (1483). 2017 Meeting Papers'.

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.

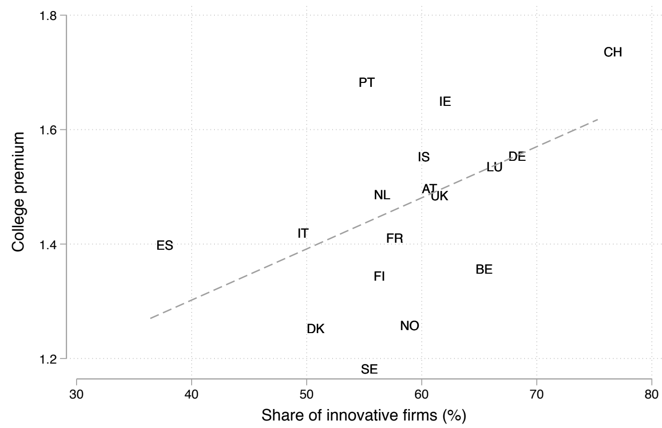
Verhoogen, E. A. (2008), 'Trade, quality upgrading, and wage inequality in the mexican manufacturing sector', The Quarterly Journal of Economics **123**(2), 489–530.

Violante, G. L. (2008), 'Skill-biased technical change', The New Palgrave Dictionary of Economics **2**.

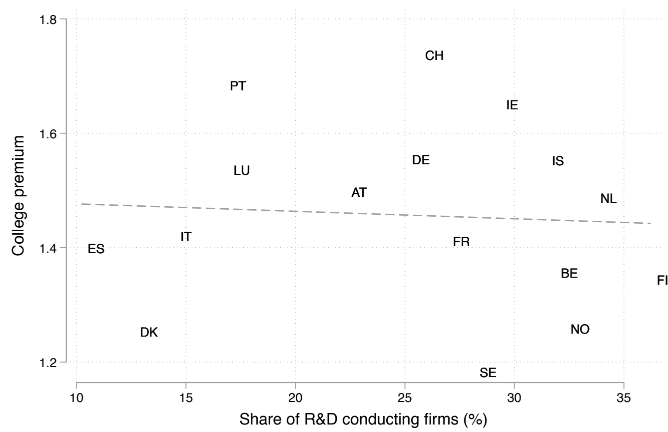
Figures

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

	Hungary		Norway	
Variable	Non-innovative	Innovative	Non-innovative	Innovative
Average age of empl.	42.1	41.3	42.2	42.5
	(0.09)	(0.10)	(5.8)	(4.7)
Share of female empl.	0.21	0.19	0.2	0.2
	(0.01)	(0.00)	(0.2)	(0.2)
Average year of education	11.4	11.8	12.8	13.6
	(0.02)	(0.03)	(1.7)	(1.7)
Share of college grad.	0.12	0.18	0.3	0.4
	(0.00)	(0.00)	(0.3)	(0.3)
Average wage (ln in NO)	173,087	206,746	7.1	7.2
	(1,672.11)	(2,446.14)	(0.6)	(0.4)
Number of employees	159	435	49.1	141.5
	(7.43)	(44.85)	(137.2)	(521.0)
ln(tangible capital/employees)	7.95	8.46	5.0	5.4
	(0.03)	(0.03)	(1.8)	(1.6)
ln(value added/employees)	8.24	8.54	6.6	6.5
	(0.01)	(0.02)	(0.8)	(0.8)

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 (2012 for Norway). 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 (Hungary)

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: Innovation and subsequent change in firm-level outcomes

Panel A: Hungary

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

Panel B: Norway

	(1)	(2)	(3)
LHS:	college wage share	college employment share	ln employment
Innovation	0.009***	0.012***	0.044***
	(0.002)	(0.002)	(0.012)
ln capital (d)	0.002	0.002	
	(0.001)	(0.001)	
ln value added (d)	-0.008***	-0.005*	
	(0.003)	(0.003)	
Dependent variable (t-1)	yes	yes	yes
Industry-year FE	yes	yes	yes
Observations	7,604	7,604	8,708
R squared	0.07	0.07	0.12

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 (11). 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)-(2) 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. Column (4) investigates the relationship between innovation and TFP change, where TFP is estimated with the methods proposed by Ackerberg et al. (2015). Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: Innovation and the college premium: worker-level regressions

Panel A: Hungary

LHS: log wage	(1)	(2)	(3)	(4)
Innovation	0.201***	0.166***	-0.028**	-0.008
	(0.022)	(0.019)	(0.012)	(0.009)
Innovation x College	0.085***	0.100***	0.123***	0.067***
	(0.027)	(0.023)	(0.014)	(0.023)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FE			yes	yes
Matched sample				yes
Observations	785,443	785,443	785,419	197,065
R-squared	0.438	0.507	0.714	0.699
Firms	6236	6236	6236	1075

Panel B: Norway

LHS: log wage	(1)	(2)	(3)	(4)
Innovation	0.107***	0.089***	-0.025***	-0.012
	(0.019)	(0.017)	(0.008)	(0.007)
Innovation x College	0.057**	0.054**	0.105***	0.040***
	(0.023)	(0.023)	(0.013)	(0.009)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FEs			yes	yes
Worker FEs				yes
Observations	8,177,865	8,177,865	8,177,865	8,137,343
R-squared	0.04	0.07	0.20	0.42
Workers	1,004,812	1,004,812	1,004,812	1,004,812
Firms	118,512	118,512	118,512	118,512
Firms in CIS	15,857	15,857	15,857	15,857

Note: This table investigates whether innovative firms pay a higher college premium with worker-level regressions, described in Equation (9), 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. 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. Column (4) includes a matched sample in Hungary and worker fixed effects in Norway. 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: incumbent workers and new entrants

Panel A: Hungary

	(1)	(2)	(3)	(4)
Innovation x New entrant	0.203***	0.176***	-0.019	-0.005
	(0.023)	(0.020)	(0.012)	(0.009)
Innovation x Incumbent	0.152***	0.143***	-0.052***	-0.016
	(0.022)	(0.019)	(0.015)	(0.013)
Innovation x College x New entrant	0.087***	0.098***	0.115***	0.044*
	(0.028)	(0.026)	(0.014)	(0.025)
Innovation x College x Incumbent	0.095***	0.097***	0.132***	0.111***
	(0.027)	(0.025)	(0.016)	(0.024)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FE			yes	yes
Matched sample				yes
Observations	785,443	785,443	785,419	197,065
R-squared	0.460	0.508	0.715	0.701
Firms	6236	6236	6236	1075

Panel B: Norway

	(1)	(2)	(3)	(4)
Innovation x New entrant	0.079***	0.084***	-0.005	-0.055***
	(0.019)	(0.018)	(0.014)	[0.013]
Innovation x Incumbent	0.100***	0.096***	-0.001	0.003
	(0.020)	(0.019)	(0.010)	(0.007)
Innovation x College x New entrant	0.067**	0.060**	0.112***	0.079***
	(0.028)	(0.028)	(0.023)	(0.016)
Innovation x College x Incumbent	0.041*	0.037	0.080***	0.020**
	(0.023)	(0.023)	(0.013)	(0.009)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FEs			yes	yes
Worker FEs				yes
Observations	8,177,865	8,177,865	8,153,963	8,137,343
Workers	1,004,812	1,004,812	1,004,812	1,004,812
Firms	118,512	118,512	118,512	118,512
Firms in CIS	15,857	15,857	15,857	15,857
R-squared	0.05	0.07	0.19	0.42

Note: This table repeats the regression in Table 4 with including interactions of the key variables with a dummy showing whether the worker had been working for at least 14 months in the firm (incumbent). 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

Panel A: Hungary

LHS: ln wage	(1) Industry-skill-year-FE	(2) Occupation year-FE	(3) One wave	(4) Two waves	(5) Pre-trends
Innovation	-0.002	-0.005	-0.007	-0.010	-0.005
	(0.009)	(0.009)	(0.007)	(0.008)	(0.010)
Innovation x College	0.074***	0.083***	0.059***	0.057**	0.069***
	(0.018)	(0.019)	(0.022)	(0.022)	(0.024)
Innovation pre-trend					0.007
					(0.013)
Innovation pre-trend x College					0.020
					(0.026)
Industry-skill-year FE	yes				
Occupation-year FE		yes			
# waves to define the inn. var	3	3	1	2	3
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes	yes
Observations	197,065	197,065	190,666	103,192	197,065
R-squared	0.700	0.752	0.699	0.686	0.699

Panel B: Norway

LHS: ln wage	(1) Industry-skill-year-FE	(2) Occupation year-FE	(3) One wave	(4) Two waves	(5) Pre-trends
Innovation	-0.010	-0.010	-0.006	-0.012	-0.010
	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
Innovation x College	0.037***	0.038***	0.014*	0.040***	0.032***
	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)
Innovation pre-trend					-0.008
					(0.006)
Innovation pre-trend x College					0.012*
					(0.006)
Industry-skill-year FE	yes				
Occupation-year FE		yes			
# waves to define the inn. var	3	3	1	2	3
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
Worker FE	yes	yes	yes	yes	yes
Observations	8,127,259	8,137,343	8,137,343	8,137,343	8,137,343
Workers	1,004,812	1,004,812	1,004,812	1,004,812	1,004,812
Firms	118,512	118,512	118,512	118,512	118,512
Firms in CIS	15,857	15,857	15,857	15,857	15,857
R2	0.42	0.43	0.42	0.42	0.42

Note: This table shows robustness checks of the worker-level regressions, described in Equation (9), with log wage as the dependent variable. 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: Innovation and the college premium: routine jobs (Hungary)

LHS: ln wage	(1)	(2)	(3)	(4)
Innovation	0.215***	0.179***	-0.022**	-0.009
	(0.018)	(0.016)	(0.010)	(0.010)
Non-routine	0.055***	0.033***	0.050***	0.060***
	(0.004)	(0.004)	(0.002)	(0.007)
Innovation x College	0.040	0.059***	0.099***	0.085***
	(0.025)	(0.022)	(0.013)	(0.021)
Innovation x Non-routine	0.050***	0.047***	0.026***	0.001
	(0.011)	(0.008)	(0.006)	(0.008)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FE			yes	yes
Matched sample				yes
Observations	784,732	784,732	784,708	157,638
R-squared	0.456	0.517	0.722	0.701
Firms	6236	6236	6236	1075

Note: This table re-runs the individual regressions in Table 4 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 8: Type of innovation and subsequent change in firm-level outcomes (Hungary)

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

Note: This table extends the firm-level regressions in Table 11 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 9: The novelty of innovation and the college premium (Hungary)

LHS: log wage	(1)	(2)	(3)	(4)	(5)
Innov x college	0.067***	0.059**	0.066***	0.062***	0.057**
	(0.023)	(0.024)	(0.023)	(0.023)	(0.025)
Innov x R&D x college		0.023			0.022
		(0.028)			(0.028)
Innov x new x college			0.004		-0.008
			(0.032)		(0.033)
Innov x developed x college				0.014	0.009
				(0.024)	(0.023)
Skill-year FE	yes	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes	yes
Observations	197,065	197,065	197,065	197,065	197,065
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 (4) of Table 4 Panel A, 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 5.9 (8.2) 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 10: The type of innovation and the college premium (Hungary)

LHS: ln wage	(1)	(2)	(3)	(4)
Tech. x college	0.069***	0.067***		
	(0.022)	(0.025)		
Org x College	0.016	0.016	0.016	0.017
	(0.022)	(0.022)	(0.022)	(0.022)
Process x college			0.048*	0.047*
			(0.026)	(0.028)
Product x college			0.038	0.037
			(0.023)	(0.023)
Innov x R&D x college		0.005		0.003
		(0.030)		(0.029)
Skill-year FE	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes
Observations	197,065	197,065	197,065	197,065
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 (4) of table 4 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 (2) 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 6.9 (1.6) percentage points higher at firms conducting only technological (organizational) innovation, while this extra premium is 8.5 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 11: How much does low and high innovation explain from the wage premium? (Hungary)

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 9 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 1: Further details of the model

The model predicts that firms invest in innovation up the point where the marginal benefit of innovation is the same as the marginal cost of innovation, formally:

$$w_Z = R\left[\lambda\left(\frac{A_{jt}}{\widetilde{A_t}}\right)(V_{jt}(\gamma A_{jt}, \kappa\theta_{jt}) - V_{jt}(A_{jt}, \theta_{jt}))\right]$$

If the benefits of innovation is larger, and so $V_{jt}(\gamma A_{jt}, \kappa\theta_{jt}) - V_{jt}(A_{jt}, \theta_{jt})$ is larger, firms are more likely to innovate. Note that these benefits can depend on the local labor market characteristics and the extent which innovation is skilled biased $\kappa > 1$.

Now let us solve the firms profit maximization problem. The profit maximization problem is the following:

$$\max_{w_{L_j}, w_{H_j}} A_{jt}\left[(1-\theta_{jt})H_{jt}(w_{H_{tj}})^{\rho} + \theta_{jt}L_{jt}(w_{L_{tj}})^{\rho}\right]^{\frac{1}{\rho}} - L_j(w_{L_j})w_{L_j} - H_j(w_{H_j})w_{H_j}$$

where $L_j(w_{L_j}) = \mathcal{L}_{L_t} w_{L_j}^{\beta_L} a_{L_j}$ and $H_j(w_{L_j}) = \mathcal{H}_{H_t} w_{H_j}^{\beta_H} a_{H_j}$.

The FOC of this problem are the following:

$$A_{jt}\left[\theta_{jt}H_{jt}(w_{H_{tj}})^{\rho} + (1-\theta_{jt})L_t^{\rho}\right]^{\frac{1}{\rho}-1} \frac{H_{jt}(w_{H_{tj}})^{\rho}}{H_{jt}(w_{H_{tj}})} \frac{\partial H_{jt}(w_{H_{tj}})}{\partial w_{H_{tj}}}\theta_{jt} = \frac{\partial H_{jt}(w_{H_{tj}})}{\partial w_{H_{tj}}} w_{H_j} + H_j(w_{H_j})$$

$$A_{jt}\left[\theta_{jt})H_{jt}(w_{H_{tj}})^{\rho} + (1-\theta_{jt})L_{jt}(w_{L_{tj}})^{\rho}\right]^{\frac{1}{\rho}-1} \frac{L_{jt}(w_{L_{jt}})^{\rho}}{L_{jt}(w_{L_{jt}})} \frac{\partial L_{jt}(w_{L_{tj}})}{\partial w_{L_{tj}}}(1-\theta_{jt}) = \frac{\partial L_{jt}(w_{H_{tj}})}{\partial w_{L_{tj}}} w_{L_j} + L_j(w_{L_j})$$

which can be rewritten to

$$A_{jt} y_{jt}^{1-\rho} H_{jt}(w_{H_{tj}})^{\rho-1}\theta_{jt}\frac{\beta_H}{\beta_H+1} = w_{H_j}$$
$$A_{jt} y_{jt}^{1-\rho} L_{jt}(w_{H_{tj}})^{\rho-1}(1-\theta_{jt})\frac{\beta_L}{\beta_L+1} = w_{L_j}$$

Taking the ratio of these two leads to the following equation:

$$\frac{w_{H_{jt}}}{w_{L_{jt}}} = \frac{\frac{\beta_H}{1+\beta_H}}{\frac{\beta_L}{1+\beta_L}} \frac{1-\theta_{jt}}{\theta_{jt}} \left(\frac{H_{jt}}{L_{jt}}\right)^{\rho-1}$$

Now take the log difference:

$$\ln \frac{w_{H_{jt}}}{w_{L_{jt}}} = \ln \frac{\beta_H}{1+\beta_H} - \ln \frac{\beta_L}{1+\beta_L} + \ln \frac{\theta_{jt}}{1-\theta_{jt}} + (\rho - 1) \ln \frac{H_{jt}}{L_{jt}}$$

Using the FOCs one can also express:

$$\ln A_{jt} + (1-\rho) \ln y_{jt} + (\rho - 1) \ln H_{jt} + \ln \theta_{jt} + \ln \frac{\beta_H}{\beta_H + 1} = \frac{1}{\beta_H} \ln H_{jt}(w_{H_{jt}}) - \frac{1}{\beta_H} \ln (\mathcal{H}_{H_t}) - \frac{1}{\beta_H} \ln a_{H_{jt}}$$

$$\frac{\beta_H}{1-\beta_H(\rho-1)} \left(\ln A_{jt} + (1-\rho) \ln y_{jt} + \ln \theta_{jt} + \ln \frac{\beta_H}{\beta_H+1} + \frac{1}{\beta_H} \ln (\mathcal{H}_{H_t}) + \frac{1}{\beta_H} \ln a_{H_{jt}} \right) = \ln H_{jt}(w_{H_{jt}})$$

The similar derivation leads the following expression for $\ln L_{jt}(w_{L_{jt}})$:

$$\frac{\beta_L}{1-\beta_L(\rho-1)} \left(\ln A_{jt} + (1-\rho) \ln y_{jt} + \ln(1-\theta_{jt}) + \ln \frac{\beta_L}{\beta_L+1} + \frac{1}{\beta_L} \ln (\mathcal{H}_{L_t}) + \frac{1}{\beta_L} \ln a_{L_{jt}} \right) = \ln L_{jt}(w_{L_{jt}})$$

Taking the difference lead to the following expression:

$$\begin{aligned} \ln \frac{H_{jt}(w_{H_{jt}})}{L_{jt}(w_{L_{jt}})} =& \frac{\beta_H}{1-\beta_H(\rho-1)} \left(\ln A_{jt} + (1-\rho) \ln y_{jt} + \ln \theta_{jt} + \frac{1}{\beta_H} \ln a_{H_{jt}} + \Lambda_H \right) + \\ & - \frac{\beta_L}{1-\beta_L\,(\rho-1)} \left(\ln A_{jt} + (1-\rho) \ln y_{jt} + \frac{1}{\beta_L} \ln a_{L_j} + \ln \theta_{jt} + \Lambda_L \right) \end{aligned}$$

where $\Lambda_H = \ln \frac{\beta_H}{\beta_H+1} + \frac{1}{\beta_H} \ln (\mathcal{H}_{H_t})$ and $\Lambda_L = \ln \frac{\beta_L}{\beta_L+1} + \frac{1}{\beta_L} \ln \left(\mathcal{H}_{L_{jt}}\right)$ are constants.

The effect of hicks neutral technological change on the skill ratio can be expressed in the following way

$$\frac{d \ln \frac{H_{jt}(w_{H_{jt}})}{L_{jt}(w_{L_{jt}})}}{d \ln A_{jt}} = \left(\frac{\beta_H}{1-\beta_H(\rho-1)} - \frac{\beta_L}{1-\beta_L\,(\rho-1)} \right) \left(1 + (1-\rho) \frac{d \ln y_{jt}}{d \ln A_{jt}} \right)$$

Note that

$$\frac{d \ln y_{jt}}{d \ln A_{jt}} = 1 + \frac{1}{\rho} \frac{1}{A_{jt}} \frac{\partial \ln \left(\theta_{jt} H_{jt}^{\rho} + (1-\theta_{jt}) L_{jt}^{\rho} \right)}{\partial A_{jt}} = 1 + \frac{1}{\rho} \frac{1}{A_{jt}} \frac{\rho \theta_{jt} H_{jt}^{\rho-1} \frac{\partial H_{jt}}{\partial A_{jt}} + \rho (1-\theta_{jt}) L_{jt}^{\rho-1} \frac{\partial L_{jt}}{\partial A_{jt}}}{\theta_{jt} H_{jt}^{\rho} + (1-\theta_{jt}) L_{jt}^{\rho}}$$

$$\frac{d\ln y_{jt}}{d\ln A_{jt}} = 1 + \frac{\theta_{jt} H_{jt}^{\rho} \frac{\partial \ln H_{jt}}{\partial \ln A_{jt}} + (1-\theta_{jt}) L_{jt}^{\rho} \frac{\partial \ln L_{jt}}{\partial \ln A_{jt}}}{\theta_{jt} H_{jt}^{\rho} + (1-\theta_{jt}) L_{jt}^{\rho}} > 0.$$

It is not hard to see that the effect of hicks neutral technological change on wage ratio is the following:

$$\frac{d\ln \frac{w_{H_{jt}}}{w_{L_{jt}}}}{d\ln A_{jt}} = -\left(\rho - 1\right) \frac{d\ln \frac{H_{jt}}{L_{jt}}}{d\ln A_{jt}}$$

Appendix 2: Country-industry level regressions

In this section we present supporting evidence at the more aggregated country-industry level.

Methodology

We start our investigations at the country-industry level to study whether broadly defined innovation is correlated with skill demand. More precisely, we look at whether the college share or the college premium increased faster in industries/countries with a higher share of innovative firms. 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.[24] These data are from Eurostat's webpage.[25]

Our question is 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.[26] In particular, we run regressions of the type:

$$\Delta y_{cst} = \beta * innovation_{cst} + \delta * y_{cst} + \eta_{ct} + \zeta_{st} + \epsilon_{cst} \tag{11}$$

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[27], the share of firms conducting R&D or the R&D intensity of the industry.[28]

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

[24] 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 reflect developments related to the Great Recession. Therefore, we stick to the 2010-2014 period.

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

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

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

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

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.

Results

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 A1 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.[29] 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 premium.

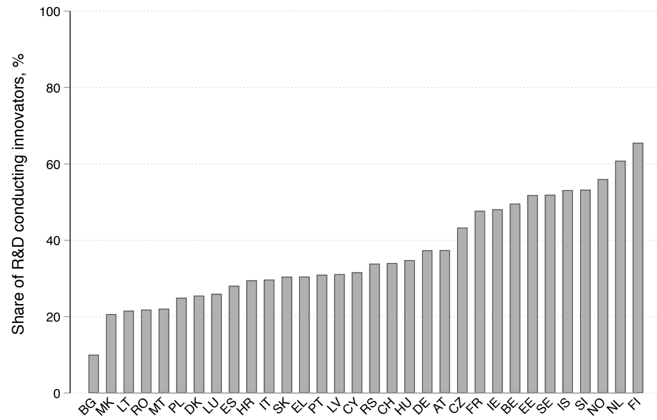
We can conclude from this exercise that broadly defined innovation is more strongly related to increasing skill demand than R&D. Low-novelty innovation, including technology adoption can be as significant in SBTC as high novelty innovation. Also, we see a response both at the quantity and the wage margin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

[29] 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.

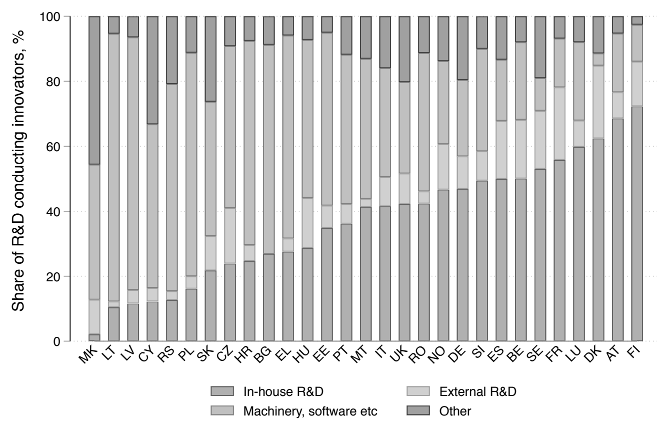
Appendix Figures

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

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



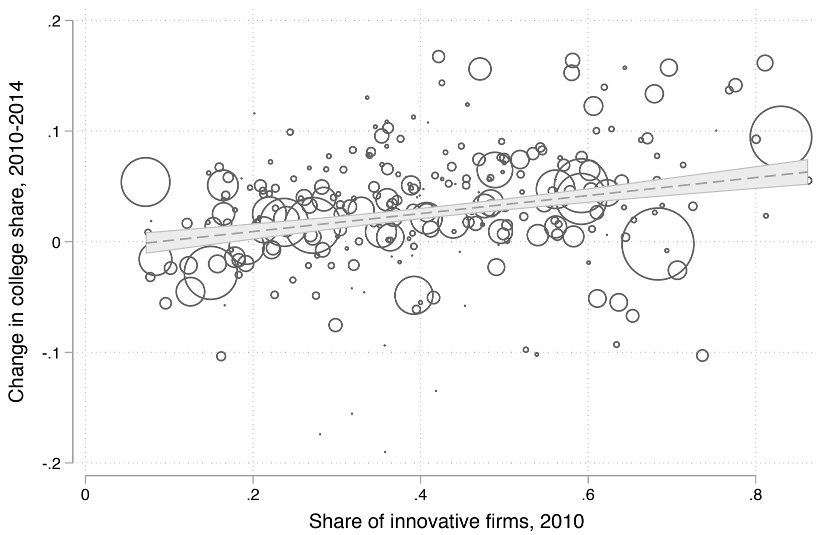
Panel B: Share of R&D in total innovation costs



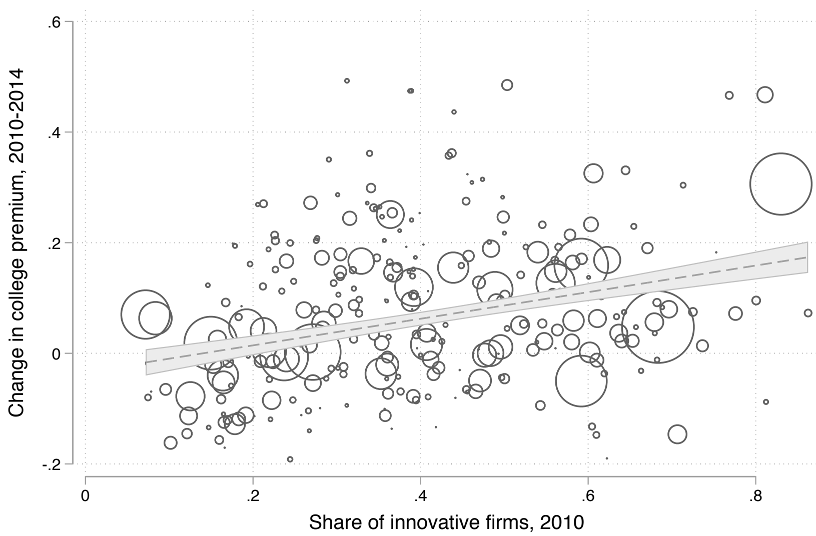
Notes: Panel A of this figure shows the share of firms which conducted R&D relative to all innovative firms by country. Panel B shows the share of different types of innovation expenditures relative to total expenditures on technological innovation. Source: CIS, 2014.

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: The share of innovative firms and growth in skill ratio, 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 show 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 A2: 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 A3: Linking the datasets: Number of firms in the sample (Hungary)

	(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 A4: Firm innovation status by industry (Hungary)

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 A5: Innovation and different worker-level outcomes (Hungary)

	(1)	(2)	(3)	(4)
LHS:	total wage	base salary	got bonus	log hours
innovation	-0.008	-0.019	0.020	-0.000
	(0.009)	(0.012)	(0.014)	(0.002)
innovation x college	0.067***	0.095***	-0.008	0.000
	(0.023)	(0.025)	(0.017)	(0.003)
Skill-year FE	yes	yes	yes	yes
Mincer variables	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes
Matched sample	yes	yes	yes	yes
Observations	197,065	197,065	197,065	197,065
R-squared	0.697	0.694	0.442	0.700

Note: This table shows the results of worker-level regressions, described in Equation (9), with different dependent variables. All the regressions follow the preferred specification from Table 4, column (4). 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 change 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 A6: Innovation and the skill premium: polarization? (Hungary)

LHS: log wage	(1)	(2)	(3)	(4)
Innovation	0.194***	0.167***	0.165***	-0.027*
	(0.035)	(0.031)	(0.030)	(0.015))
Innovation x Primary	-0.015	-0.031	-0.030	-0.009
	(0.035)	(0.030)	(0.030)	(0.011)
Innovation x Vocational	0.022	0.018	0.017	0.006
	(0.027)	(0.025)	(0.025)	(0.008)
Innovation x College	0.092***	0.101***	0.086***	0.114***
	(0.032)	(0.026)	(0.026)	(0.013)
Skill-year FE	yes	yes	yes	yes
Mincer variables		yes	yes	yes
Firm FE			yes	yes
Matched sample				yes
Observations	785,443	785,443	785,419	197,065
R-squared	0.438	0.507	0.714	0.699
Firms	6236	6236	6236	1075

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 4 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: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.