Firm-level Technological Change and Skill Demand *

Attila Lindner
UCL, IFS, CRES

Balázs Muraközy
University of Liverpool

Balázs Reizer
CERS

Ragnhild Schreiner
University of Oslo

March 2021

Abstract

We propose a novel approach to quantify the contribution of firm-level technological change to skill demand and aggregate inequality. The basic idea behind our approach is that firms' wage and employment decisions reveal information about the nature of technological change when there is imperfect competition in the labor market. Skill-biased technological change increases *both* the firm-level skill ratio and the skill premium. In contrast, other shocks, including firm-specific output demand shocks or labor supply shocks, have opposite effects on the skill ratio and the skill premium, and so they cannot explain an increase in both outcomes. We apply this idea by exploiting administrative data from Hungary and Norway linked to the Community Innovation Surveys (CIS), which uniquely provide a direct measure for a broad class of technological changes at the firm-level. We show that—even conditional on workers' observable and unobservable skills—firms increase both their college wage premium and college–non-college ratio following technological change. Our estimates imply that technological change taking place over a 10-year period increased the aggregate college premium by 6.1% in Norway and by 13.8% in Hungary. In line with the two countries' different distance from the technological frontier, this was mainly driven by R&D-based and high novelty innovation in Norway and mostly by technology adoption in Hungary. These results highlight that technological change is still a key driver of aggregate trends in inequality, even if the college premium has been falling recently.

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

JEL-codes: J31, J24, O30, O33

*We thank Dávid László, Hajnalka Katona, Dániel Pass and Jon Piqueras for excellent research assistance. We thank Ufuk Akcigit, Richard Blundell, Sydnee Caldwell, Márton Csillag, Christian Dustmann, Jan Eeckhout, Scott Imberman, Pat Kline, Miklós Koren, Magne Mogstad, John van Reenen, Kjell Salvenes, Uta Schoenberg, Kjetil Storesletten and seminar participants at Central European University, London School of Economics, University College London, Queen Mary University, IAB, the University of Oslo, the Norwegian School of Economics and Statistics Norway for their insightful comments. This research has been funded by the European Union's Horizon 2020 research and innovation programme (grant agreement number 822390 for Murakozy and grant agreement number 949995 for Lindner). Lindner acknowledges financial support from the Economic and Social Research Council (new investigator grant, ES/T008474/1.)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. This research is part of Oslo Fiscal Studies (OFS) at the Department of Economics, University of Oslo. OFS is supported by the Research Council of Norway, project number 267428. Finally, we are grateful to the Central Statistical Office of Hungary and Statistics Norway for making the data available.

1 Introduction

Technological change is the main driver of economic growth. Nevertheless, the gains from new ideas may disproportionately benefit high skilled workers such that technological progress increases income inequality (Acemoglu 2002, Goldin & Katz 2010). In this paper, we study the consequences of technological change on inequality by focusing on the role of firms. Firms play a crucial role in the diffusion of new technologies through the process of innovation (Mokyr 2003, Bloom et al. 2016). At the same time, a growing number of studies document that firms contribute to rising aggregate inequality (Song et al. 2015, Card et al. 2018). As a result, it is important to understand how firm-level innovation activities and inequality are interlinked.

Direct evidence on the impact of technological change on inequality is still scarce, and existing evidence is somewhat inconclusive.[1] Moreover, our understanding of the relationship between innovation and inequality is nearly exclusively based on easily measurable proxies of innovation, such as R&D and patents, which are unlikely to capture a large part of firm- or economy-level technological change. For instance, in France, one of the more innovative countries in Europe, only 39% of innovative firms conducted R&D continuously, 27% occasionally and 34% reported innovations without any R&D spending. Internal R&D constituted 55% of French firms' innovative expenditures, while the remaining 45% was spent on external R&D, machinery, software and other investments.[2] On the output side, only 12% of French innovators applied for a patent. The role of R&D in innovation inputs and the share of innovators applying for patents are substantially lower in less innovative countries, where technology adoption plays a larger role, such as Hungary (Appendix Figure B.3).This paper instead measures technological change with a broad definition of firm-level innovation activities that involve the introduction of production processes, products or management methods that are new to the firm, but not necessarily new to the market or the world. Among the European countries this innovation measure strongly correlates with country-level college premium (see Figure 1).[3]

Understanding the contribution of technological change to the college premium is especially relevant in light of the recent decline in college-to-non-college wage premium observed in many developed countries. For instance, between 2000 and 2015, the college premium decreased by 5 percentage points in the United States, by 11 percentage points in Norway and by 15 percentage points in Hungary.[4] The drop in college premium might reflect that technological change, which had been favoring college educated workers from the 80s to the early 2000s (Katz & Murphy 1992), altered its character and favors other groups in the economy now. Nevertheless, this recent fall in aggregate college premium has coincided with a significant higher education expansion in all three countries, which may mask a substantial contribution of technology to inequality.

[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] These numbers refer to in-house R&D in firms introducing product and/or process innovations.

[3] In Appendix Section A.1 we provide further details about this relationship and show that the positive correlation is robust to controlling for the share of R&D conducting firms, the college ratio or GDP/capita. In Appendix Section A.2 we also provide some additional evidence by exploiting country-industry level variation in innovation activities.

[4] The country-level college premia and the college ratios come from the OECD Education at a Glance 2014 and 2020 data. The college premium refers to the wage difference between college workers and workers with higher secondary degrees. For Norway the college premium is missing for 2000 and so we report the changes between 2005 and 2015.

To better understand the contribution of technological change to these aggregate trends, we examine first the relationship between firm-level technological change and skill demand. Motivated by our empirical finding that firm-level innovation leads to an increase in the skill premium, we deviate from the canonical models in the literature of technological change, and introduce imperfect competition in the labor markets. We follow Card et al. (2018) and Manning (2013) and assume that firms do not take wages as exogenously given, but rather they need to set higher wages in order to expand.

The main insight from our model is that in the absence of skill-biased technological change, a negative relationship emerges between the skill ratio and the skill premium at the firm level. Intuitively, the "law of demand" implies that when the relative price of an input goes up, relative demand for that input falls. This logic holds even if both the skill premium and skill ratio are endogenously determined in the model. Then we show that skill-biased technological change can increase both the skill premium and skill ratio. In contrast, other type of (confounding) shocks that potentially coincide with firm-level innovation (e.g. firm-specific output demand shocks, labor supply shocks) either increase the college ratio and decrease the college premium or *vice versa*, but they cannot explain an increase in both outcomes.

This insight is similar to the one provided by Katz & Murphy (1992) in relation to the U.S. wage structure over the 80s, where they argued that a positive relationship between relative skill prices and quantities suggests that the technological change is skill-biased.[5] In this paper, we highlight that the same reasoning can be made about firm-level relative supply and demand when there is imperfect competition in the labor markets.

It is worth emphasizing that the identification assumptions required to assess the extent to which technological change is skill-biased are weaker than what is needed to identify the impact of innovation on firm-level productivity. A key concern for identifying the latter is that innovative firms might foresee, and start innovating in response to, some positive demand shocks, and this can contaminate estimates of innovation on overall productivity. However, such a firm-level shock will not interfere with our proposed test for skill-biased technological change. This is because only relative input demands, and not the level of output, matter for assessing the skill-bias. Identification, in fact, mainly relies on two standard assumptions often made in the literature. First, we assume that there is a constant elasticity of substitution between high skilled and low skilled workers (standard CES production function), and so simply changing the level of production does not alter the marginal rate of transformation, the skill ratio and the skill premium. Second, we assume that firms optimize both before and after innovation such that the first order conditions from the firm's problem hold in each period. In that sense, identifying the extent to which technological change is skill-biased is less challenging than assessing the impact on firm-level productivity, which could be biased by these shocks.

Guided by our model, we investigate empirically whether innovation activities lead to an increase in the skill premium and the skill ratio at the firm level. We use exceptionally rich micro data from two

[5] As we described above, the skill premium does not increase in more recent data even though the skill ratio continued to increase, and so the aggregate evidence is less conclusive.

countries, Norway and Hungary, that are at very different distances from the technological frontier. In Norway, R&D based, high novelty innovation dominates while in Hungary relatively few firms innovate and if so, they often adopt technologies developed elsewhere. This allows us to compare two very different innovation systems. In both countries, we have access to the rich information available from the European Community Innovation Survey (CIS), which allows us to identify firm-level technological changes in a comprehensive way.

We estimate the change in skill premium by implementing a difference-in-differences type identification strategy where we compare changes in the premium of workers at firms which start to innovate to changes in the premium in firms not changing their innovation status. We find that innovation is associated with a 2-4 percent increase in the wage premium in Norway, and a 5-6 percent increase in Hungary. This increase in skill premium is permanent and present even 5 years after innovation, is not driven by higher bonus payments, and arise for both new entrants and incumbent workers. We also find that the increase in the skill premium emerges after innovation and is not driven by pre-innovation wage premium differences. Finally, to ensure that the increase in college premium is not driven by simply the compositional change of the workforce, we also control for unobserved workers skills by exploiting our particularly rich data from Norway, where we can follow workers across firms.

Our estimates of the skill premium are robust to including a variety of controls for market-specific shocks, that could potentially be correlated with firm-level innovation. In particular, we include local labor market-specific time trends, industry-skill-group-specific time trends and occupation-specific time trends in our robustness tests. Further, the estimates are not sensitive to alternative timing assumptions, and also robust to allowing for unobserved heterogeneity in firm-specific college premiums.

To assess the impact of firm-level innovation on the skill ratio, we implement a similar difference-in-differences identification strategy. In particular, we estimate how innovation is related to subsequent long (six-year) changes in the skill ratio at the firm level. Estimating a long difference is suitable for capturing the long-term effects of innovation, while at the same time adjusting for unobserved (time invariant) firm heterogeneity. This strategy closely follows Caroli & Van Reenen (2001), who study the effect of innovation on skill demand in French and British firms. In line with their findings, we find that innovation is associated with subsequent growth in the college share.

We also demonstrate that, in line with the predictions of the imperfect labor market competition model presented here, we find more muted wage responses and a larger increase in the skill ratio in local areas with higher firm density, where firms have more limited wage-setting power.[6] Finally, we also corroborate our key findings by documenting the changes in skill demand in response to innovation activities induced by a quasi-exogenous change in an R&D tax credit policy in Norway.

These findings highlight that technological change tends to be skill-biased both in Norway and Hungary. Using our estimates in the change in skill demand, we also quantify the contribution of technological progress to the change in the aggregate college premium. Firm-level innovation activities

[6] In our model there is a tight link between local area firm density and firms' wage-setting power. This comes from the observation that firms' wage-setting power depends on the dispersion of workers' idiosyncratic preferences for working at particular firms, and this dispersion is likely to be larger if commuting time between firms is larger.

contribute to aggregate inequality through two channels. First, reallocation of skilled workers to innovative firms, which pay higher wages. Second, our estimates of the wage premium suggest that firms pay higher premium to workers following innovation. Using our estimates of the change in college ratio and the college premium following innovation, we calculate that skill-biased innovation contributes by 6.1 percentage points to the increase in the aggregate skill premium in Norway and by 13.8 percentage points in Hungary over ten years.

Finally, we assess whether there is heterogeneity in the contribution of different types of innovation to inequality. A common pattern in both countries is that both innovation with technical aspects (product or process innovation) and organizational changes are skill biased. Nevertheless, the bulk of the contribution to aggregate inequality comes from firms combining technical with organizational changes. At the same time, we find a difference between Norway and Hungary with respect to R&D and high-novelty innovation. In Norway, firms conducting R&D-based and high-novelty innovations are responsible for the majority of the changes in skill demand. In contrast, non-R&D and low-novelty innovations, which are associated with technology adoption, play a key role in Hungary. This latter finding underscores that technological change is skill-biased even in countries farther away from the technology frontier.

Our paper is related to several strands of the literature. First, we contribute to the large literature that relates the evolution of wage inequality to skill-biased technological change (see, for example Autor et al. 1998, Acemoglu 2002, Goldin & Katz 2010, Acemoglu & Autor 2011a). Instead of inferring the change in skill bias from aggregate trends in relative skill ratio and skill premium, we exploit the fact that most technologies diffuse slowly and firms play a crucial role in this process (Griliches 1957). By focusing on firm-level changes in technology and applying a difference-in-differences strategy we can net out the effect of changes in institutions (Bound & Jonson 1992, DiNardo et al. 1996, Stansbury & Summers 2020) and market power (De Loecker et al. 2020), and focus on solely the contribution of technological change. Our strategy also differs from Haanwinckel (2018) who, similarly to us, recognizes the crucial role of firms, but instead of directly studying changes in skill demand at the firm level, he builds a model of tasks within firms and infers technological change from the aggregate changes in worker-firm sorting and in the distribution of firm-level skill premia.[7]

Our paper also contributes to the literature that directly considers technological change (or innovation) and firm-level skill demand. Many papers in the literature focus on specific technologies, such as computers (see, e.g. Krueger 1993, DiNardo & Pischke 1997, Dunne et al. 2004, Beaudry et al. 2010), broadband internet (e.g. Akerman et al. 2015, Hjort & Poulsen 2019), robots (e.g. Graetz & Michaels 2018), artificial intelligence (e.g. Frank et al. 2019), automation (Doms et al. 1997, Acemoglu et al. 2020) or high-novelty innovation, such as R&D (Bøler 2015, Aghion et al. 2017) and patents (Kline et al. 2019). In this paper, we consider a much wider range of innovation activities that is likely to capture most forms of technological change taking place in the economy, including adoption of technologies by firms far from the technology frontier. Furthermore, we take a step further and also quantify the contribution of firm-level technological changes to aggregate inequality.

[7] Haanwinckel (2018) introduces imperfect competition on the labor market into a task-based framework, while here we apply the standard CES production function. In principle it is possible to derive estimable reduced form equations between changes in task content and firm-level technological change, but such an analysis is beyond the scope of our paper. Nevertheless, in Section 5 we empirically assess the change in task content.

Focusing on a wider range of innovation activities is not unprecedented in the literature (Caroli & Van Reenen 2001, Bresnahan et al. 2002, Abowd et al. 2007). Nevertheless, these studies usually rely on relatively small cross-sectional surveys that measure specific innovation activities. In contrast, our data includes five repeated waves of a large-scale innovation survey, where each wave covers a large number of firms (around 5,000), and provides consistent measures for various types of innovation activities over time (and across countries). The richness of the data allows us to better control for the fact that innovating and non-innovating firms tend to be different. The panel dimension of our survey also allows us to account for compositional changes following innovation, which leaves us with more credible estimates of the effect of innovation on the skill premium. Finally, our paper also makes a methodological advancement relative to these papers by highlighting the issues of simply focusing on the skill ratio to assess the skill-biasedness of technological change. The changes in skill ratio can be confounded by shocks to labor supply or even by output demand shocks if firms' wage-setting power differs between low and high skilled workers. Naturally, instrumental variable strategies can alleviate these issues, but appropriate instruments may be hard to find.[8] Our approach instead studies the changes in skill premium as well, which even though requires data on wages, it is often more readily available in many contexts.

Our paper also relates to a growing number of papers studying responses to firm-level shocks with imperfect competition in the labor markets (e.g. Card et al. 2018, Garin & Silvério 2018, Lamadon et al. 2018, Kroft et al. 2020). The fact that we find an increase in the firm-level skill premium following innovation is consistent with some wage-setting power of firms. The implied firm-specific labor supply elasticity is between 2-3, which is consistent with recent quasi-experimental estimates from the literature (e.g. Caldwell & Oehlsen 2018, Cho 2018, Kroft et al. 2020, Dube et al. 2017, Bassier et al. 2020). We also demonstrate that, consistent with the prediction of the model, the implied firm-specific elasticity is tightly linked to the local labor market firm density. In denser areas, firms have a more limited wage-setting power as the implied firm-specific labor supply elasticity is around 4, while in areas with very low density firms operate almost like a local monopsony, with a firm-specific labor supply elasticity is less than one. These geographic differences also suggest that technological change can affect differently rural and urban labor markets.

We also contribute to the literature about the heterogeneity of innovation. One strand of this literature quantifies and compares innovation with technological aspects and organizational changes. The seminal paper of Caroli & Van Reenen (2001) shows that both types of innovations are skill biased, while Evangelista & Vezzani (2010) focuses on productivity and shows that firms which conduct a broader range of innovation activities—for example, combine technological with organizational innovation—have a higher performance. Another dimension, the distinction between R&D and non-R&D innovation, was emphasised by Lopez-Rodriguez & Martinez-Lopez (2017), who show that non-R&D innovation also contributes to productivity. Our contribution is that we compare the skill bias of all these different types of innovation and quantify their aggregate effect on the skill premium. Our results show that all these different types of innovation are skill-biased to a certain extent, but

[8] In fact, even estimates identified from an exogenous change in technology can lead to spurious findings of skill biased technological change if firms' wage-setting power differs between skilled and unskilled workers. In that case, the exogenous increase in Hicks-neutral productivity can increase the skill ratio even if there is no skill bias (see Proposition 1).

their absolute and relative contribution depends on the context.

In what follows, Section 2 outlines the relationship between technological change, skill demand and relative wages of skilled and non-skilled workers when there is imperfect competition in the labor markets. Section 3 describes our data sources and the institutional context in Norway and Hungary. Section 4 discusses our empirical strategy to estimate the change in skill ratio and skill premium following innovation. The results of these estimations are presented in Section 5, while we quantify the aggregate implication of changes in firm-level skill demand in Section 6. Finally, Section 7 concludes.

2 Conceptual framework

We study the impact of firm-level technological change on the skill premium and the skill ratio. Motivated by our empirical findings showing that firm-level technological change has an impact on firm-level wages, we endow firms with some wage-setting power. This wage-setting power arises from worker heterogeneity in their valuation of jobs due to non-wage related characteristics, as in Card et al. (2018).

We start by describing the firm's problem and then examine how firm-level technological change affects employment and wages. We assume that there are J firms, each of them using two inputs in production at time t: high-skilled labor (H_{jt}) and low-skilled labor (L_{jt}).[9] In the empirical part we will proxy skills by education and, therefore, throughout the paper we use skills and education interchangeably.

Firms produce output (Q_{jt}) with the following CES technology in every period:

$$Q_{jt} = A_{jt}\left[\theta_{jt}H_{jt}^{\frac{\sigma-1}{\sigma}} + (1-\theta_{jt})L_{jt}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \sigma \geq 0 \tag{1}$$

where A_{jt} is the Hicks-neutral productivity term, while θ_{jt} is the skill bias productivity term measuring the extent to which the technology used by the firm is skill-biased.[10] Importantly, technological change affects one or both of these productivity terms. Extending this production function with capital or other intermediate inputs is a relatively straightforward exercise as we demonstrate in Appendix C.[11]

Following Violante (2008), we define skill-biased technological change as an increase in the

[9] While in our conceptual framework we abstract away from worker's heterogeneity within a skill group, such heterogenity can be incorporated into our framework. The presence of worker heterogeneity would complicate the discussion, while our results would hold after netting out changes in worker's composition. We carefully deal with worker heterogeneity within a skill group in our empirical implementation though. See more details in Section 4.

[10] Given the wide range of technological change our measures capture, we remain agnostic about what exact mechanism drives skill bias. The increase in skill demand can come from capital-skill complementary (see e.g. Krusell et al. 2000), from better ability of skilled workers to deal with new technologies (see e.g. Nelson & Phelps 1966), or from "flatter" organizations (see e.g. Milgrom & Roberts 1990).

[11] We add capital by applying a nested CES structure. However, the results can be generalized to any production function of the following structure: $F(Q_{jt}, K_{jt})$, where Q_{jt} comes from Equation (1) and K_{jt} denotes capital. Note that such a production function rules out that capital is more complementary to high-skilled than to low-skilled workers (see e.g. Krusell et al. 2000). We consider such complementarity between capital and skills as one formalization of skill-biased technological change (Violante 2008), which we approximate with a change in θ_{jt}.

marginal rate of transformation (MRT) between skilled and unskilled workers. In our production function an increase in θ will always increase the MRT, therefore it represents skill-biased technological change.[12]

Firms maximize their profit given this production function:

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

and they face the following budget constraints:

$$Q_{jt} = A_{jt} \left[\theta_{jt} H_{jt})^{\frac{\sigma-1}{\sigma}} + (1-\theta_{jt}) L_{jt})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{3a}$$

$$\ln p_{jt} = \frac{1}{\rho} \ln \kappa_{jt} - \frac{1}{\rho} \ln Q_{jt} + \frac{\rho-1}{\rho} \ln p_t + \frac{1}{\rho} \ln I_t \tag{3b}$$

$$\ln L_{jt}(w_{Ljt}) = \ln(L_t \Lambda_{Lt}) + \beta \ln w_{Ljt} + \ln a_{Ljt} \tag{3c}$$

$$\ln H_{jt}(w_{Hjt}) = \ln(H_t \Lambda_{Ht}) + \beta \ln w_{Hjt} + \ln a_{Hjt}, \tag{3d}$$

Budget constraint (3a) just restates the production function defined above. Budget constraint (3b) represents a downward sloping output demand function that can be micro founded using a monopolistic competition framework (see Appendix C). In this constraint, p_{jt} is the price of the firm's product, ρ is the elasticity of demand, κ_{jt} captures firm-specific demand shifters, p_t denotes the price index in firm j's market at time t, while I_t is the income spent on total consumption in firm j's market in period t.[13]

The third (3c) and fourth (3d) budget constraints represent the upward sloping labor supply functions firms face. These firm-level labor supply curves can be micro founded in a discrete choice framework as follows (Card et al. 2018). 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

$$u_{iSjt} = \ln(\tau w_{Sjt}^{\lambda}) + \ln a_{Sjt} + \epsilon_{iSjt}, \tag{4}$$

where τ and λ approximate the income tax system (see Lamadon et al. 2018), $\ln a_{Sjt}$ is a firm-specific amenity that is common to all workers in group S, while ϵ_{iSjt} captures idiosyncratic preferences of worker i for working at firm j, arising from commuting distance, work flexibility and so on. We assume that the ϵ_{iSjt} are independent draws from a type-I Extreme Value distribution with a dispersion parameter ϕ. As demonstrated by Card et al. (2018), under these assumptions, the approximate

[12] An alternative way to write the production function is as follows:

$$Q_{jt} = \left[(A_{Hjt} H_{jt})^{\frac{\sigma-1}{\sigma}} + (A_{Ljt} L_{jt})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

In our case, $A_{jt} = \left(A_{Hjt}^{\frac{\sigma-1}{\sigma}} + A_{Ljt}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$ and $\theta = (A_{Hjt}/A_{jt})^{\frac{\sigma-1}{\sigma}}$. Note that an increase in A_{Hjt}/A_{Ljt} in this formulation only favors skilled workers if $\sigma > 1$. When $\sigma < 1$, then a decrease in A_{Hjt}/A_{Ljt} leads to skill-biased technological change. Our θ will increase in both of these cases.

[13] When different firms serve different markets, p_t and I_t are market-specific. However, to make the notation simpler, we suppress this index in our derivations.

firm-specific upward-sloping labor supply functions lead to Equations (3c) and (3d), where $\beta = \lambda/\phi$ and $\ln(L_t \Lambda_{Lt})$, $\ln(H_t \Lambda_{Ht})$ represent local labor market conditions. An important implication is that the firm-specific labor supply elasticity is decreasing in the dispersion of worker preferences, because the more heterogeneous workers are, the more firms need to raise the wage to attract workers. A special case of this model is the perfectly competitive labor market, where the dispersion of workers' idiosyncratic preferences converges to zero—meaning that all workplaces are homogeneous from the workers' perspective. In this case, β is infinite and so firms face a perfectly elastic labor supply function.

The first order conditions from the firm's profit maximization problem lead to the following relationship between the relative wages and skill ratio at the firm level:

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

where $\frac{\theta_{jt}}{1-\theta_{jt}}$ measures the extent to which technology is tilted toward high-skilled labor. This equation resembles the key equation describing the relationship between relative demand and relative wages of college and non-college workers in the skill-biased technological change literature (see e.g. Katz & Murphy 1992, Violante 2008, Goldin & Katz 2010). One important difference, however, is that in our case such a relationship emerges at the firm level, with firm-specific skill premia and ratios.

Since $\sigma > 0$, Equation (5) highlights that relative wages and relative demand are negatively related if the technological change is not skilled-biased (θ_{jt} is unchanged). Intuitively, the negative relationship is driven by a firm-level "law of demand": if the relative price of an input increases, firms will substitute away from that input. Even though firm-level relative wages and employment both change endogenously in a setting where labor markets are non-competitive, they still remain negatively related in the absence of skill-biased change. Consequently, if we observe that relative skill ratios and relative wages are both positively affected by technological change, we can infer that the technological change is skill-biased.

The derived Equation (5) together with budget constraints in Equations (3a)-(3d) imply that

$$\ln \frac{w_{Hjt}}{w_{Ljt}} = \frac{\sigma}{\sigma+\beta} \ln \frac{\theta_{jt}}{1-\theta_{jt}} - \frac{1}{\sigma+\beta} \ln \frac{H_t \Lambda_{Ht}}{L_t \Lambda_{Lt}} - \frac{1}{\sigma+\beta} \ln \frac{a_{Hjt}}{a_{Ljt}} \tag{6a}$$

$$\ln \frac{H_{jt}}{L_{jt}} = \frac{\beta\sigma}{\sigma+\beta} \ln \frac{\theta_{jt}}{1-\theta_{jt}} + \frac{\sigma}{\sigma+\beta} \ln \frac{H_t \Lambda_{Ht}}{L_t \Lambda_{Lt}} + \frac{\sigma}{\sigma+\beta} \ln \frac{a_{Hjt}}{a_{Ljt}}, \tag{6b}$$

These equations highlight that the relative skill and wage ratios do not depend on the Hicks-neutral part of the production function (A_{jt}) or the various (firm-specific) output demand shocks (e.g. κ_{jt}). Instead, those relative terms will depend on the extent to which the technology relies on skilled workers (θ_{jt}), on the relative firm-level amenities (a_{Hjt}/a_{Ljt}), and on the market-level labor supply shocks in the two markets ($H_t \Lambda_{Ht}/L_t \Lambda_{Lt}$).

Indeed, the changes of the skill premium and the skill share depend on A_{jt} and θ_{jt} the following way:

$$\underbrace{\Delta \ln \frac{w_{Hjt}}{w_{Ljt}}}_{\substack{\text{Change in}\\ \text{skill premium}}} = \frac{\sigma}{\sigma+\beta} \underbrace{\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}}_{\substack{\text{Change in}\\ \text{skill bias}}} - \frac{1}{\sigma+\beta} \underbrace{\Delta \ln \frac{H_t \Lambda_{Ht}}{L_t \Lambda_{Lt}}}_{\substack{\text{Change in}\\ \text{market-level}\\ \text{labor supply}}} - \frac{1}{\sigma+\beta} \underbrace{\Delta \ln \frac{a_{Hjt}}{a_{Ljt}}}_{\substack{\text{Change in}\\ \text{relative}\\ \text{amenities}}} \tag{7a}$$

$$\underbrace{\Delta \ln \frac{H_{jt}}{L_{jt}}}_{\substack{\text{Change in}\\ \text{skill ratio}}} = \frac{\beta\sigma}{\sigma+\beta} \underbrace{\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}}_{\substack{\text{Change in}\\ \text{skill bias}}} + \frac{\sigma}{\sigma+\beta} \underbrace{\Delta \ln \frac{H_t \Lambda_{Ht}}{L_t \Lambda_{Lt}}}_{\substack{\text{Change in}\\ \text{market-level}\\ \text{labor supply}}} + \frac{\sigma}{\sigma+\beta} \underbrace{\Delta \ln \frac{a_{Hjt}}{a_{Ljt}}}_{\substack{\text{Change in}\\ \text{relative}\\ \text{amenities}}}, \tag{7b}$$

where Δ denotes the change between before and after innovation.

These equations motivate our difference-in-differences style regressions described in detail in Section 4. We will study the change in the skill premium and also the change in the skill ratio following innovation and compare it to the change at non-innovative firms. According to these equations, skill-biased innovation—an increase in θ—will affect positively (or non-negatively, if $\beta = \infty$) both the skill ratio and the skill premium. At the same time, other firm- or market-specific shocks either have no effect on the skill ratio (e.g. the Hicks-neutral increase in the production function, A_{jt}, or the change in output demand, κ_{jt}) or have an opposite effect on the skill ratio and the skill premium (e.g. the relative change in labor supply or the relative change in amenities).

It is worth emphasizing that even if technological change is initiated as a response to some firm-specific demand shock (e.g. a change in κ_{jt}), those shocks will not affect the skill ratio and skill premium as they do not play a role in Equations (7a) and (7b). It is well known that separating the effect of innovation on TFP (or A_{jt}) from output demand shocks that coincide with innovation is quite challenging (see e.g. Crépon et al. 1998, Griffith et al. 2006). Nevertheless, we do not need this separation here. Even if innovation responds endogenously to output demand shocks, our strategy remains valid since those shocks do not directly affect relative wages and the skill ratio. Still, to make sure that our results do not simply reflect the endogenous nature of firm-level innovation, we present evidence from Norway where we exploit an exogenous change in incentives to invest in innovation (see Section 5.3).

The equations also highlight the key advantage of identifying skill-biased technological change from the changes in both the skill ratio and in the skill premium. When identifying solely from the change in skill ratio, dealing with the relative changes in labor supply and the relative change in amenities is crucial.[14] For instance, the increase in the skill ratio might simply reflect that firms invest more in innovation when they expect a change in their relative supply or relative amenities. If someone focuses solely on the skill ratio then it is important to directly control for many hard-to-observe factors or to exploit changes in innovation activities that are orthogonal to potential confounders. In fact, as we will see later, even if someone has access to fully exogenous technology shifters, focusing solely on

[14] As we discuss later (see Proposition 1), whenever the firm-specific labor supply elasticities differ by skill groups, even firm-specific demand shocks can generate an increase in skill ratio, but they cannot explain an increase in both the skill ratio and the skill premium.

skill ratio can be problematic whenever the firm-specific labor supply elasticities differ by skill groups (e.g. $\beta_H \neq \beta_L$). In that case even Hicks-neutral shocks can increase the skill ratio, but such shocks cannot increase both the skill ratio and skill premium (see Proposition 1). Our approach instead relies on the observation that market-level (relative) labor supply shocks ($H_t \Lambda_{Ht}/L_t \Lambda_{Lt}$) or changes in relative amenities (a_{Hjt}/a_{Ljt}) will either increase the skill ratio and decrease the skill premium or *vice versa*, but they do not lead to a simultaneous increase in both outcomes. Three points should be noted.

First, the above derivation assumes that workers' qualities are constant within skill groups. Yet, a potential reason why relative wages change after innovation is that firms may hire higher quality workers within a given skill group. One can even imagine that innovation only affects worker sorting to firms, and not the wage premium paid to equally productive workers. As a result, it is crucial to control for firm-level changes in worker composition following innovation. To deal with this, we will show that changes in the wage premium are present also for incumbent workers who had worked at the firm before innovation (as well as new entrants). We also exploit our particularly rich data from Norway which allow us to control for worker's unobservable characteristics. We discuss in detail how we deal with that issue in Section 4.

Second, as we have mentioned above, competitive labor markets are a special case of the model where β is infinite. In this case firm-level innovation should only affect the skill ratio, and not the skill premium. Furthermore, changes in firm-level skill premium should reveal no information about the firm's technology. Nevertheless, the one-equation strategy of focusing only on the skill ratio could still be problematic. For instance, if firms are more likely to innovate in markets where the skill premium is declining and the researcher cannot control for that, then the increasing skill ratio will simply reflect the change in relative prices and not the effect of innovation. Nevertheless, documenting that there is no change (or an increase) in the firm's skill premium can still be used to rule out that such confounders drive the change in the skill ratio.[15]

Third, Equations (7a) and (7b) highlight that if the firm-specific labor supply is more elastic (β is higher), then the responsiveness to changes in the skill bias (θ) gets smaller for the skill premium and larger for the skill ratio. Remember, the firm-specific labor supply elasticity, β, is related to the dispersion in the idiosyncratic preferences of individuals working at a particular firm (see Equation (4) and the subsequent discussion). This dispersion is likely to be related to the average distance between various workplaces within the labor market due to differences in commuting time, and so firm density à la Ciccone & Hall (1996) is a good proxy for such dispersion. In line with this prediction, we show in Section 5 that in local areas with high firm density the increase in the skill premium is smaller, while the increase in the skill ratio is larger following innovation.

[15] Caroli & Van Reenen (2001) propose to apply a one-equation empirical strategy. They assume a translog production function, which is a second order approximation of our CES production function around $\sigma = 1$. Whenever $\sigma = 1$ (Cobb-Douglas production function), the change in the wage share, which we can be calculated by adding up Equations (7a) and (7b), will be equal to the change in skill bias:

$$\Delta \text{wage share of } \mathrm{H}_{jt} = \Delta \ln \frac{w_{H_{jt}}}{w_{L_{jt}}} + \Delta \ln \frac{H_{jt}}{L_{jt}} = \Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}$$

Nevertheless, the change in share of high-skilled wages does not capture the change in skill bias when $\sigma \neq 1$.

Extensions. We make two extensions of the basic set-up that we discuss here briefly and in more detail in Appendix C and Appendix D. So far we have assumed that the labor supply elasticities of low and high skilled workers are similar. In our framework this comes from the assumption that dispersion of the idiosyncratic error term in Equation (4) is not skill specific. While this is a common assumption in the literature, we relax this assumption in Appendix C. Suppose that budget constraints (3c) and (3d) are replaced with the following:

$$\ln L_{jt}(w_{Ljt}) = \ln(L_t \Lambda_{Lt}) + \beta_L \ln w_{Ljt} + \ln a_{Ljt} \tag{3c$'$}$$

$$\ln H_{jt}(w_{Hjt}) = \ln(H_t \Lambda_{Ht}) + \beta_H \ln w_{Hjt} + \ln a_{Hjt}, \tag{3d$'$}$$

where β_L and β_H refer to the firm-level labor supply elasticities of low and high skilled workers, respectively.

When the firm-level labor supply elasticities differ, we cannot express the skill premium and the skill ratio in a closed form. Furthermore, it can be shown that even Hicks-neutral productivity shocks can affect both the skill premium and the skill ratio. Nevertheless, it is possible to prove the following proposition.

Proposition 1. *Suppose that firms maximize profit given the budget constraints in Equations (3a), (3b), (3c′), (3d′). Changes in A_{jt} and κ_{jt} have the following effect on firm-level skill ratio $\left(\ln \frac{H_{jt}}{L_{jt}}\right)$ and skill premium $\left(\ln \frac{w_{Hjt}}{w_{Ljt}}\right)$.*

1. *If $\beta_H = \beta_L$, then $\ln \frac{w_{Hjt}}{w_{Ljt}}$ and $\ln \frac{H_{jt}}{L_{jt}}$ are unaffected by A_{jt} and κ_{jt}.*
2. *If $\beta_H > \beta_L$, then $\ln \frac{w_{Hjt}}{w_{Ljt}}$ is decreasing and $\ln \frac{H_{jt}}{L_{jt}}$ is increasing in A_{jt} and in κ_{jt}.*
3. *If $\beta_H < \beta_L$, then $\ln \frac{w_{Hjt}}{w_{Ljt}}$ is increasing and $\ln \frac{H_{jt}}{L_{jt}}$ is decreasing in A_j and in κ_{jt}.*

Proof. See Appendix Section C.2. □

Proposition 1 states that Hicks-neutral changes (A_{jt}) and firm specific demand shifters (κ_{jt}) directly affect the skill ratio and the skill premium if the dispersion of idiosyncratic preferences differs across the two skill groups ($\beta_H \neq \beta_L$). Nevertheless, the effects of these shocks on $\ln \frac{w_{Hjt}}{w_{Ljt}}$ and $\ln \frac{H_{jt}}{L_{jt}}$ always have opposite signs. So if one of them increases, the other will fall. This implies that demand shifters (κ_j) or Hicks-neutral shocks (A_j) cannot explain a joint increase of the skill premium and the skill ratio.

Why does even a Hicks-neutral change (A_{jt}) affect the skill ratio when $\beta_H \neq \beta_L$? When a firm experiences an increase in A_{jt}, it will expand and, therefore, increase its demand for both types of workers. If, for example, $\beta_H > \beta_L$, high skilled workers are more responsive to changes in wages than the low skilled ones, and, therefore, firms can expand their skilled labor force more when they increase their wages. In optimum, firms adjust both on the wage and quantity margins: they raise high skilled workers' wages less $\left(\Delta \ln \frac{w_{Hj}}{w_{Lj}} < 0\right)$, but hire more of them $\left(\Delta \ln \frac{H_j}{L_j} > 0\right)$.

An important implication of Proposition 1 is that documenting an increase in skill ratio after innovation does not imply that innovation has a skill-biased productivity term, even if the innovation shock *per se* was truly exogenous. In the presence of imperfect competition in the labor market, even a Hicks-neutral change in the production function can affect the skill ratio (for instance, if $\beta_H > \beta_L$). Nevertheless, as Equation (5) above demonstrated, whenever both the skill premium and the skill ratio increase, technological change must be skill-biased.

Labor market power. We discuss another extension of the standard framework. So far we have assumed that agents are atomistic in labor markets, and so they do not take into account how their behavior affects other agents' behavior. We relax this assumption and incorporate strategic interactions into our framework by following Berger et al. (2019a) and Deb et al. (2020). In particular, Deb et al. (2020) show that Equation (5) has to be extended with an extra term capturing the change in market power in the presence of strategic interactions in the labor market (see more details in Appendix D)[16]:

$$\underbrace{\Delta \ln \frac{w_{Hjt}}{w_{Ljt}}}_{\substack{\text{Change in}\\ \text{skill premium}}} = \underbrace{\Delta \ln \frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}}_{\substack{\text{Change in}\\ \text{markdown}}} + \underbrace{\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}}_{\substack{\text{Change in}\\ \text{skill bias}}} - \frac{1}{\sigma} \underbrace{\Delta \ln \frac{H_{jt}}{L_{jt}}}_{\substack{\text{Change in}\\ \text{skill ratio}}} \tag{8}$$

where $\Delta \ln \frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$ shows the change in relative firm specific markdowns for firm j operating in labor market m. In the presence of strategic interactions, the mark-downs are firm-specific and depend on the firm's market share of the particular skill group in the local labor market (Berger et al. 2019a). These market shares may themselves be affected both by skill-biased and Hicks-neutral technological change. Importantly, if Hicks-neutral innovation leads to a large increase in market shares (and so an increase in market power), that can introduce a positive correlation between the change in the college premium and the college ratio. To investigate this possibility, we estimate the change in market shares and relative markdowns following innovation and use the above Equation (8) to separate the change in skill bias from the change in labor market power. We find some evidence for change in relative market power following innovation in Norway, but the overall impact of that on skill-premium is limited (more details on this can be found in Appendix D).

3 Data and Institutional Set-up in Hungary and Norway

In our empirical application, we study the contribution of innovation activities and technological change to skill demand in two European countries. We provide a basic overview of the two countries here and further details can be found in Appendix B. Norway is one of the richest and most developed countries in the world, with a GDP/capita level which is 20% larger than that of the USA. Hungary is among the poorest European Union member states, with a GDP/capita slightly above 50% of USA level in PPP terms. In terms of innovation activities, Norway is classified as a "Strong innovator" (similar to France, ranked 10th in the EU out of 28) while Hungary as a "Moderate innovator" (ranked

[16]This extension explicitly models labor markets, and, therefore, we index the labor market-level variables with m.

23rd) according to the European Innovation Scoreboard.[17] This suggests that Norway is much closer to the technology frontier and its innovation activities often expand the world technology frontier, while Hungarian firms rely more heavily on technology adoption to move closer to that frontier.

Labor market institutions also differ between Norway and Hungary. Norway's labor market is an example of the Nordic model, which has three key features: (i) flexible hiring and firing, (ii) a generous social safety net and (iii) active labor market policies. Union density is very high, with more than 35% of workers in the private sector being Union members in 2012. Collective bargaining with the participation of Unions has led to smaller wage dispersion and sustained high wage growth. Centralized collective bargaining determines a floor for wage levels and increases, but there is considerable scope for deviations from these industry-level agreements. Indeed, firm-level wage agreements often lead to substantially higher wages, allowing for firm-level wage setting. For the majority of white-collar workers in the private sector, centrally negotiated collective agreements do not specify wages, and therefore these workers have only firm-level wage formation with strong individual-level elements (Nergaard 2014).

Hungarian employment protection institutions are closer to the Anglo-Saxon institutions than to those found in most continental countries. It is relatively easy to dismiss workers (Tonin et al. 2009) and wage bargaining takes place mostly at the individual level. Union membership is very low and coverage of collective industry-level agreements is limited and usually lax rules are set.

Providing evidence from these two countries that differ in various important dimensions such as distance to technological frontier and strength of labor market institutions allows us to draw a more complete picture of the impact of technological change on skill demand, compared to what is possible by studying just one country. The similarities of our findings in the two countries across many dimensions suggest that our results have some external validity to other countries and contexts as well.

Now we describe the key data sets used for the analysis, namely the innovation survey and the employer-employee data, which we are able to link for each country.

3.1 Innovation data (CIS)

The first data source is the Hungarian and Norwegian versions of the Community Innovation Survey (CIS), conducted in a harmonized way in the European Union member states and some other countries, including Norway. The richness of the CIS has been 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 to the best of our knowledge, no paper has used so far the CIS to assess the relationship between innovation and skill demand.

The survey is bi-annual and covers a representative sample of manufacturing and service firms

[17] https://ec.europa.eu/growth/industry/policy/innovation/scoreboards_en. This ranking is multi-dimensional, based strongly on the CIS. Norway is not an EU member state, but its score can be compared to member states' score. Based on data from 2018.

in the economy. The survey ask about innovation activities in the survey year and the preceding two years: for example the CIS 2014 refers to innovation activities in 2012, 2013 and 2014. In this paper we use six waves of the CIS survey from the period between 2004 and 2014 (five waves for Norway: 2004-2012). In both countries, the sample size has been progressively increasing from about 4,000 firms in 2004 to more than 7,000 at the end of the period of study.

The main advantage of this survey is that it provides direct, reliable and broad measures for innovative activities of the firm. As a result, the innovation measured by the CIS is a good measure for all the different types of firm-level technological change. Consequently, throughout the paper we use technological change and innovation interchangeably.

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).[18] Consequently, innovation in the CIS is defined very broadly, namely, as the introduction of products, services, processes and organizational solutions which are new or significantly modified from the viewpoint of the firm, but that are not necessarily new to the market.

The survey allows us to investigate the heterogeneity of innovation along a number of dimensions. First, it distinguishes between different types of innovation which can be classified into two main categories: innovation with technical aspects (product and/or process innovation), and organizational innovation.[19] Second, the survey allows us to look at heterogeneity based on whether the innovation is R&D-based, and in terms of its novelty. For the empirical implementation, we create an R&D dummy that takes the value one if the firm reports positive in-house R&D spending, and consider an innovation *new* if the firm reports the innovation to be new to its market.

3.2 Matched employer-employee data

We link the CIS data to employer-employee data from both countries.

3.2.1 Norway: Employer-employee register

The employer-employee register,[20] provided by Statistics Norway, contains all employment spells and annual information on wages and days worked per employment spell. We merge the employer-employee

[18] 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.

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

[20] A more detailed description of this database is available at https://www.nav.no/en/home/employers/nav-state-register-of-employers-and-employees.

register to data on worker demographics that include information on level of education, age and gender. Finally, we link these data to the balance sheet of limited liability companies.

To study the impact of innovation on skill premium 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 in each year. We restrict the sample to those aged 19 to 67. To be included in the data we further require that the worker is employed in a firm for at least 30 days in a given year. This gives an unbalanced panel data set containing 8,330,444 observations with 1,013,857 workers employed in 118,967 different firms over the 12-year period 2002-2013. This data set is merged to five waves of the CIS survey for Norway that was conducted biannually from 2004-2012, and covers the years 2002-2012. This gives an unbalanced sample consisting of 4,804,373 worker-year observations in 15,530 unique firms. To study firm-level changes in skill ratio, we aggregate up from the worker-level sample.[21]

3.2.2 Hungary: Structure of Earnings Survey

In Hungary, we use the Structure of Earnings Survey (Bértarifa) database, which is a survey harmonized across EU countries.[22] This is an annual worker-level survey, which includes information on a number of demographic variables, including schooling, job characteristics, tenure and on the wage that 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, especially if they move across firms.

These data can be merged to administrative balance sheet data collected for tax purposes by the National Tax and Customs Administration (NAV). This database includes employment, industry classification and balance sheet information of all double-entry bookkeeping enterprises in Hungary.

When creating our worker-level regression file, we start out with the Structure of Earnings Survey for the years between 2003-2014. The 12 waves of the survey consist 2,085,455 individuals and 42,395 unique firms. We merged the survey to 6 waves of the CIS which was conducted biannually between 2004 and 2014. The merged sample consists 785,443 individuals and 6,236 unique firms.[23] For the firm-level regressions we need to observe the firm and its workers both at the year of the innovation survey and 6 years afterwards (see Section 4.2). Therefore, when creating the sample for the firm-level regressions, we can only use firms which were surveyed in the CIS waves in 2004, 2006 or 2008, and were in the Structure of Earning Survey both in that year and 6 years afterwards. These restrictions reduce the firm-level regression sample to 2363 firm-year observations and 1733 unique firms.

[21] In the firm-level analysis, part-time workers are included. Since hourly wages cannot be reliably calculated for part-time workers, we drop them from the worker-level analysis.

[22] More information about this survey is available at https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Structure_of_earnings_survey_(SES).

[23] After matching firms based on observable characteristics we have 179,065 worker-year observations and 1,716 unique firms (see Section 4.1).

3.3 Descriptive Statistics

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 (see e.g. Griffith et al. 2006), innovative firms are larger in both countries. 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. Innovative firms also pay substantially higher wages. In terms of age composition, the innovative and non-innovative firms are very similar in both countries. Finally, the differences in data sources in the two countries imply that in Norway we cover more small firms than in Hungary.

4 Empirical Approach

Guided by our model, the main aim of the empirical analysis is to estimate the extent to which innovation or firm-level technological change is tilted toward skilled workers. To do so, we will assess the change on firm-level skill ratio and skill premium following innovation.

4.1 Estimating the Change in Skill Premium

To estimate the relationship between innovation and the college premium, we start from a Mincer-type wage regression. In particular, our benchmark empirical model is the following:

$$\ln wage_{ijt} = \delta^u innov_{jt} + \delta^s innov_{jt} \times college_i + \gamma X_{ijt} + \eta_i + \varphi_j + \varsigma_{kt} + \varepsilon_{ijt}, \tag{9}$$

where $wage_{ijt}$ is individual i's wage at firm j at time t. $college_i$ is an indicator variable taking the value one if worker i has college education, and $innov_{jt}$ is an indicator variable taking the value one if the firm innovates in the current or any of the previous two CIS waves.[24] The vector X_{ijt} contains Mincer-type control variables, including gender, age, tenure, tenure squared, an indicator variable taking the value one if the worker is a new entrant to the firm and education dummies (including $college_i$) in specifications without worker fixed effects.[25] In the benchmark specification we also include worker fixed effects (η_i, only in Norway), firm fixed effects (φ_j), and various group-specific time effects denoted by ς_{kt} in the equation above. In the benchmark specification ς_{kt} includes (1-digit) industry-time fixed effects and 4 education group-time effects. By including the interacted education group-year effects we effectively control for education-specific wage trends, as well as policy changes that might affect education groups differently, such as changes in the minimum wage. In a more saturated model we also include industry-location-year fixed effects, occupation-location-year fixed effects or industry-occupation-location-year fixed effects.

In the regressions above δ^s, the coefficient on the interaction between $college_i$ and $innov_{jt}$,

[24]The CIS survey is a biannual survey and therefore, if a firm reports innovation in CIS wave at time t, then we set $innov_{jt}$ equal to one between $t-1$ and $t+4$ according to this definition.

[25]For Hungary, we additionally include controls for hours worked and a dummy for part-time employees as part-time employees are included in the sample.

captures the the change in skill premium following technological change. As our conceptual framework demonstrated (see Equation (7a)), the change in skill premium, δ^s, depends on the change in skill-biasedness, θ, and the change in market-level labor supply.[26]

To filter out market-level labor supply shocks, we explore multiple control groups for innovative firms by including various industry-year, location-year, and occupation-year fixed effects (ς_{kt}) in the regression.[27] These controls filter out the changes in skill premium that arise at the labor market level. Nevertheless, controlling for market-level changes at the very detailed level can be problematic if there are spillovers from innovative (treated) to non-innovative (untreated) firms within a narrowly defined market. Such spillovers would lead to the violation of Stable Unit Treatment Value Assumption (SUTVA) and bias our estimates. Nevertheless, our results are not sensitive to including even very detailed location-occupation-industry-year fixed effects, which suggests that the bias caused by such spillover effects must be limited in our context.

However, even after controlling for market-level labor supply, the endogenous nature of innovation could potentially lead to spurious findings. Including person effects (η_i) and firm effects (φ_j) in Norway goes a long way to control for unobserved differences in wage-setting policies and workforce quality between innovative and non-innovative firms. One remaining concern is that firms with better management may pay a higher wage premium even before innovation and may also be more likely to innovate. While we find no indication for such wage premium differences in the data when we study pre-trends, we also control for unobserved time-invariant differences in skill premium in some regressions. Furthermore, we also present evidence based on an exogenous shift in innovation activities induced by a change in R&D tax credits in Section 5.3.

In Hungary, given that we cannot follow workers across firms, we instead estimate our main empirical specification in Equation (9) on a matched sample where innovative firms are compared to non-innovative ones with similar pre-innovation characteristics.[28] This matching procedure together with firm fixed effects in the regression allow us to alleviate the concern of endogeneity coming from the inherent differences in pay structure that might be present even before innovation. Again, the lack of pre-trends suggests that the matching procedure handles any pre-existing differences in firm-level skill premium.

Another important concern with interpreting δ^s as an estimate of the change in skill premium is that innovation might lead to a change in the composition of the workforce. If higher productivity

[26]The skill premium might also be affected by changes in firm-specific amenities. Nevertheless, if someone is solely interested in understanding the contribution of technological change to relative wages (and not relative utilities), it does not matter whether the impact on wages goes through disproportionately lowering amenities of high-skilled workers or through tilting the production function toward high-skilled workers. This might explain why most papers in the rent-sharing literature rule out the possibility that firm-level shocks affect workers' amenities (see e.g. Kline et al. 2019, Lamadon et al. 2018).

[27]We classify industries based on 1-digit European Industry-standard classification system (NACE codes) rev.2, and occupations based on 2-digits ISCO 08 codes (International Standard Classification of Occupations).

[28]We describe in detail the matching procedure in Appendix B. In a nutshell, our matching procedure works as follows. For "treated" firms that are not innovative in the first time we observe them in the CIS, but become innovative in a later wave, we match them with firms that never innovate in any CIS wave. We use firm-level characteristics for matching at the time of the first appearance in the CIS survey when, by our sample restriction, neither the treatment nor the control innovate. We use the following variables from the balance sheet for matching: 1-digit industry dummies, year dummies, log employment, log productivity, log wage premium and ownership. We also add variables from the CIS following Griffith et al. (2006) and match based on the main market of the firm, the types of funding it received.

college educated workers self-select to more innovative firms, then the estimated change in the skill premium will simply reflect the change in quality of the workforce and not a genuine increase in wages. In the benchmark specification we include various measures of worker characteristics to filter out potential compositional changes in terms of observables (gender, age, tenure, tenure squared). We also estimate the impact of the policy on a sample of incumbents, namely workers that had already been working in the firm prior to innovation and so are unaffected by the changes in composition. Including person effects (η_i) in our benchmark regression for Norway also helps to deal with unobserved differences in the quality of workers.

Furthermore, since firms might pay heterogeneous skill premia to their workers, not taking this into account could potentially lead to a bias in worker effects, which could contaminate our control for worker quality in Norway. Therefore, we also present robustness to the inclusion of firm specific skill premium in the regressions for Norway. Nevertheless, to avoid controlling for too many fixed effects in the regressions, and so by throwing out the identifying variation, we group firms into deciles based on their college premium and then we include an additional interaction of firm premium-type deciles with the college dummy in the regression Equation (9). We also explore an iterative procedure for classifying firms as described in Appendix Section B.4.

4.2 Estimating the Change in Skill Ratio

In order to estimate how innovation or technological change is related to subsequent change in the skill ratio, we start from Equation (7b). As suggested by that equation, we use a difference-in-differences type of strategy, where we compare firms which innovated at the beginning of the period with non-innovators in the same industry with similar initial characteristics. In particular, we follow Caroli & Van Reenen (2001) and estimate long-difference regressions of the form:

$$\Delta y_{jt} = \delta innov_{jt} + \gamma \Delta X_{jt} + \gamma^y y_{jt-1} + \varsigma_{kt} + \epsilon_{jt} \tag{10}$$

On the left-hand side, we have the change between year t and $t+6$ at firm j in outcome y (such as share of college workers, college to non-college ratio). $innov_{jt}$ is the same as the key variable in the worker regressions, i.e. an indicator variable taking the value one if the firm innovates in the current or previous two CIS waves. If innovation is skill-biased, it will lead to an increase in θ_{jt}, and, therefore to a positive estimated δ. Following Caroli & Van Reenen (2001), we control for the changes in the firm's capital and value added, denoted by ΔX_{jt}.[29] The specification differences out time invariant firm and labor market characteristics, while the industry-year fixed effects (ς_{kt}) aim to capture industry-level labor supply shocks $\left(\Delta \ln \frac{H_t \Lambda_H}{L_t \Lambda_L}\right)$. Finally, controlling for a lagged value of the outcome variable (y_{jt-1}) captures initial firm heterogeneity. Standard errors are clustered at the firm level. As argued by Caroli & Van Reenen (2001), such a long difference specification is likely to capture the long-run effects of innovation, as opposed to short-run fluctuations in outcomes.

[29] The regression sample is reduced to firms in the CIS waves conducted up to and including 2008, since we cannot observe long-term outcomes for firms innovating after 2008. We winsorize all the long difference variables at the 5th and 95th percentiles.

5 Empirical Results

5.1 Innovation and Changes in Skill-demand

Skill premium. We start our analysis by studying the relationship between innovation and the skill premium. Table 2 shows the estimates from the benchmark specification (Equation (9)) for Norway (Panel A) and for Hungary (panel B). Column (1) shows results on the full sample when only skill-year fixed effects are included, which control for labor market level shocks. According to the results for Norway, workers without a college degree earn 10.5 (s.e. 1.9) percent more in innovative firms (relative to workers with similar education levels in non-innovative firms), while this difference is 18.2 (10.5 + 7.7) percent for college educated workers (compared to college educated workers in non-innovative firms). The cross sectional wage premium of innovative firms is somewhat larger in Hungary, with low- and high-skilled workers earning 20.1 (s.e. 2.2) and 28.6 (20.1 + 8.5) percent more in innovative firms. In column (2) we also control for worker observable characteristics such as age, tenure and tenure square, which do not explain much of the innovative firm wage premium. In column (3), we further include firm fixed effects. In this specification, the low-skilled innovation premium becomes negative in both countries, while the college innovation premium becomes even higher than before, at 13 (s.e. 1.6) and 12.3 (s.e. 1.4) percent 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 workers.

Our benchmark 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 (as described in Section 4). Importantly, while the estimates become smaller in both countries, they remain highly significant both in economic and statistical terms. In Norway, high-skilled employees experience a 4.5 (s.e. 1.0) percent wage increase following a successful innovation, while this effect is 6.7 (s.e. 2.3) percent in Hungary. Overall, we find remarkably similar results in the two countries, with a 4-7 percent increase in college educated workers' wage premium following innovation.

In columns (5) and (6), we include one and two innovation pre-trend dummy variables indicating that the firm will innovate in the subsequent CIS wave, or in the CIS wave after the following one, as well as the interaction of these pre-trends with the college dummy. We do not find any evidence of a pre-trend in skill premium in any of the two countries, underscoring that the main results do not reflect pre-existing wage premium differences between innovative and non-innovative firms.

A number of additional robustness checks are presented in Table 3, all starting from our preferred specification (column (4) of Table 2).

Filtering-out market-level labor supply shocks. Our first concern is that innovation may be correlated with market-level labor supply shocks. As we have described in Section 2, these shocks have an opposite impact on the skill ratio and the skill premium and so they should not explain an increase in both outcomes. Still, it is worth exploring how sensitive our skill premium estimates are to various ways of controlling for market-level shocks. Columns (1)-(5) of Table 3 control for labor market-time specific shocks based on various definition of labor markets. Column (1) includes (1-digit) industry-year fixed effects, as well as district-year fixed effects. These additional controls are included to control for time-varying product or labor market specific shocks. In column (2), we include industry-district-year fixed effects to control for industry-specific time effects within local labor markets. Column (3) includes (2-digit) occupation-district-year fixed effects to additionally take out local-level occupation-specific shocks.[30] In column (4), we include industry-occupation-district-year fixed effects. This latter specification takes out shocks occurring at narrowly defined labor markets. Nevertheless, focusing on very narrow labor markets might also be problematic as innovative firms' decisions may affect non-innovative ones. For instance, if some innovative firms hire skilled workers and pay higher wages, then non-innovative firms will also need to pay higher wages. Such spillover effects will lead to a downward bias in our estimates. In line with that, by comparing Column (4) with our benchmark specification, we find that controlling for these fixed effects reduces the college premium estimates by around 20% (e.g. 4.5%, s.e. 1.0, in benchmark vs. 3.6%, s.e. 1.1, with industry-occupation-district-year fixed effects for Norway). Still, overall, our estimates are quite robust to applying labor market controls defined at various levels and in each specification we find that the college premium increases significantly following innovation both in Norway and Hungary.

In column (5) of Table 3, we also explore the possibility that the impact of local labor market shocks varies by firm type. For instance, in Proposition 2 in Appendix C we show that if $\beta_H \neq \beta_L$, then the very same labor supply shock might have a differential impact on firms operating in the same labor market, depending on the skill bias term of the firm (θ_{jt}). To deal with this issue, we classify firms into quartiles based on their initial skill ratio (which is a monotonous function of the unobserved θ_{jt}, according to Equation (6b)) and we include quartile-district-year fixed effects in the regression.[31] The estimated change in the college premium is still substantial (2.2%, s.e. 1.1, for Norway and 5.7%, s.e. 1.9, for Hungary) and statistically significant, suggesting that the observed increase in the wage premium cannot be attributed to a change in market-level wage index (Λ_{Ht}) or supply of skills (H_t).

Short and medium-term effects. In columns (6) and (7) of Table 3, we investigate the shorter term impact of innovation. Recall that in our benchmark specification we examine the average change in the wage premium up to 7 years after innovation. In column (6), we study the impact of innovation up to 3 years after innovation,[32] while in column (7) up to 5 years. In both countries, the effect of innovation increases with time, which is in line with innovation having a gradual effect on firms – also confirming that our estimate is not simply driven by short-term changes in college premium resulting from a temporarily higher effort of implementing technological change. Nevertheless, even when looking at the immediate effects of innovation, we find a clear increase in the college premium.

Controlling for firm-specific college premium. As we have described in Section 4.1, a potential concern with regression Equation (9) is that person effects will be biased in the presence of heterogeneous firm-level skill premium (pre-innovation differences in θ_{jt}). Column (8) shows

[30] For Norway, the data on occupation comes from Statistics Norway's statistics on monthly earnings.

[31] We classify firms into skill ratio quartiles based on their initial year in our sample, which is the starting year of our analysis (2002 for Norway and 2000 for Hungary) or the entry date if the firm enters later.

[32] In particular, if a firm report that it has been implemented an innovation activity in the CIS survey, then $innov_{jt}$ will equals to one between in the years $t-1$ and t in regression equation (9). Note that in some cases firms will innovate in t and in some cases it will do it $t-1$ and our survey does not allow to distinguish that. As a result, this definition is conservative and we expect to underestimate the true impact of innovation on wage premium.

the estimates when we control for this heterogeneity by including firm-level skill premium deciles and worker effects in the regression. The estimates in column (8) are very similar to the baseline estimates (3.5% vs. 4.5% in the baseline). This highlights that even after allowing for variation in the firm-specific wage premium, we find a clear and significant increase in the college premium following innovation.

New entrants vs. Incumbents. In Table 4 we explore whether the change in wage premium differs for incumbent and new entrant workers. A key characteristic of our framework is that the increase in the wage premium results from firms having to pay higher wages when hiring new workers following the innovation. This implies that, in contrast to some rent sharing models where incumbent workers obtain some rent following firm level shocks (such as in e.g. Kline et al. 2019), new hires should also receive a higher skill premium following innovation. Table 4 reports results for incumbents and new entrants. In particular, we interact the incumbent and new entrant dummies with the *innovation* dummy and the interaction of *innovation* $\times$ *college*, with the triple interactions showing the effect of innovation on incumbents and new entrant college workers.[33] The results show that both new entrants and incumbent high-skilled workers receive a higher premium following an innovation in both countries. In both countries new entrants experience a larger increase in wages than incumbents (4.2% vs. 2.0% in Norway and 9.1% vs. 3.8% in Hungary). These results are in line with the monopsonistic wage setting applied in our model, where firms need to raise wages both for new entrants and incumbents if they want to attract more workers.

Effect on structure of worker's compensation and hours worked. In Appendix Table A.3, we also explore whether innovation has an impact on the structure of workers' remuneration. Our baseline hourly wage measure includes base salary and all other financial compensation (e.g. overtime, bonuses). Nevertheless, our data allow us to look at the impact of innovation on various components of earnings separately. We find that the increase in the base salary after innovation is very similar to the increase in total salary. This confirms that the changes in the skill premium following innovation (presented in Table 2) are not driven by increases in bonus payments rewarding successful innovation. We also find no indication for changes in hours worked. Finally, we also show that even if we include non-cash benefits in worker's compensation in Norway where this information is available, the estimated change in college premium is very similar to the benchmark specification.

Polarization. So far we have classified workers into two skill groups, and looked at whether innovation affects the skill premium of college relative to non-college workers. However, Acemoglu & Autor (2011b) argue that the middle-skilled occupation categories, such as middle-skilled clerical, administrative, production and operative occupations, tend to be more affected by technological change than either high- or low-skilled occupation categories, and that this has contributed to the observed wage polarization in the US. We explore the presence of wage polarization by interacting the innovation dummy in Equation (9) with the four-category schooling variable (primary schooling,

[33] In Norway we define incumbents as working at the firm for at least 6 years, and new entrants are all other workers. This definition ensures that all incumbent workers were at the firm before innovation started. If we use an alternative definition where we define incumbents as working at the firm for at least 2 years, we get a very similar estimate. In Hungary, our data structure does not allow to define incumbents as working for at least 6 years at the company. As a result, there we look at the short term impact of innovation (column (6) in Table 3) and define incumbents as working at the firm for at least 24 months.

secondary schooling, vocational education, and college, see Appendix Table A.4). In Norway, workers with vocational training earn a wage premium following innovation relative to workers with only primary or secondary education (see Column 4 of Table A.4). In Hungary, in contrast, the wages of the lower three educational categories do not seem to change after innovation takes place, while the wages of college educated workers increase substantially. Therefore we find little support for any negative impact of innovation on middle education groups—if anything, there is some increase in wage premium for that group in Norway.

We also explore whether the effect of innovation differs between routine and non-routine occupations (see e.g. Autor et al. 2003). In Appendix Table A.5 we estimate regression Equation (9) by including routine intensity and its interaction with the innovation dummy. The results from this exercise are presented in Table A.5. We find that people working in less routine jobs are paid higher wages in general, but there is no unusual increase in their premium following innovation. At the same time, the college premium increases by a similar magnitude in these regressions, which suggests that the increase in skill demand following innovation is not limited to non-routine occupations.

Change in the skill ratio. We assess the impact of technological change on the skill ratio by estimating regression Equation (10). The main results are presented in Table 5 for the two countries. Column (1) shows the impact of innovation on the long difference of the share of college educated workers in total employment. We find a significant positive relationship in both countries: the college employment share increases by 1.1 (s.e. 0.2) percentage points in Norway and by 1.9 (s.e. 0.8) percentage points in Hungary during the six-year period following firm innovation. These estimates are very close to what Caroli & Van Reenen (2001) found for British and French firms. Column (2) shows that the college to non-college ratio increases by around 2.8 (s.e. 0.6) percentage points in Norway, which is a 5.7 (s.e. 1.2) percent increase in skill ratio from the non-innovative firms' average college ratio (0.49, see Table 1). For Hungary, we find that the skill ratio increases by 2.9 (s.e. 0.8) percentage points subsequent to innovation activities, which is equal to a 14 (s.e. 4) percent increase from the non-innovative firms' average college ratio (0.2 see Table 1). Column (3) shows that innovation is associated with stronger employment growth, with a significant estimate in Norway. The main takeaway from the firm-level results is that innovation leads to an increase in the share of high-skilled workers, though the increase in skill ratio (at least when measured in percent change) is more prominent in Hungary than in Norway.

5.2 Implications

So far we have documented that there is a clear increase both in the skill premium and the skill ratio subsequent to innovation activities. As we have described in detail in Section 2, an increase in both of these outcomes provides *prima facie* evidence for firm-level technological change being skill-biased. In our model, Hicks-neutral technological changes or firm-level output demand shifters do not affect the skill ratio (whenever $\beta_H = \beta_L$), while market-level labor supply shocks or changes in amenities can explain an increase in one or the other, but not in both. Furthermore, as we demonstrated in Proposition 1 (and Proposition 2 in Appendix C), even if $\beta_H \neq \beta_L$, it is hard to reconcile the joint increase in the skill ratio and the skill premium with Hicks-neutral technological changes, firm-level output demand shifters or market-level labor supply shocks.

Equations (7a) and (7b) also highlight that the extent to which firms respond to skill-biased technological change (increase in θ) via the skill premium versus the skill ratio margin depends on the elasticity of labor supply (β). In particular, the impact of $\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}$ on the skill premium is $frac\sigma\sigma + \beta$, while its impact on the skill ratio $\frac{\beta\sigma}{\sigma+\beta}$. This has two implications.

First, the ratio of the impact on the skill ratio relative to the impact on the skill premium is roughly equal to β, the elasticity of firm-level labor supply. For Norway, the estimated increase in the skill premium varies between 4.5% (column 4 in Table 2) and 2.2% (column 5 in Table 3). The estimated increase in the skill ratio is 5.9%. Consequently, the implied firm-level labor supply elasticity is between 1.3 and 2.7. For Hungary, the estimated increase in the skill premium varies between 6.9% (column 5 in Table 2) and 5.5% (column 4 in Table 3), while the change in the skill ratio is 14%. The implied firm-level labor supply elasticity, therefore, is between 2 and 2.6. These estimates are remarkably similar to each other and are also in the range of the existing estimates in the literature. For instance, Saez et al. (2019), studying pay-roll tax cuts in Sweden, find that the elasticity of firm-specific labor supply is between 1.8 to 2.4.[34]

Second, whenever firms face a more elastic labor supply, we expect a relatively large response on the quantity margin: a relatively larger impact on the skill ratio and a smaller impact on the skill premium. Remember that, in our model, the firm-level labor supply elasticity, β, is a function of ϕ, the dispersion of workers' idiosyncratic preferences for working at a particular firm. A key component of this dispersion is commuting distance, which is presumably smaller in local areas with higher firm density (or areas where the average distance between firms are smaller). Consequently, we expect that in local areas with a high firm density (and a low average commuting time), firms face a more elastic labor supply and therefore the increase in the skill ratio will be larger, while the increase in the skill premium will be smaller.

In Figure 2 we explore heterogeneity in the post-innovation change of the skill ratio and the skill premium by the spatial density of firms. Like in Ciccone & Hall (1996), we measure firm density as the average number of firms per square kilometer in the the local area. Then we estimate whether the changes in the skill ratio and the skill premium depend on firm density. We describe the estimation strategy in more detail in Appendix Section A.5, while in Figure 2 we show the estimated change in the skill ratio and skill premium at the 10th percentile (blue bar) and at the 90th percentile (gray bar) of the across-firm distribution of spatial density. Reassuringly, we find in both countries that in high density areas the increase in the skill ratio is larger, while the increase in the skill premium is smaller.

The estimates in Figure 2 show that in the lowest density areas on Norway, the change in the skill ratio is 0.6% (s.e. 2.2%), while the skill premium increases by 7% (s.e. 1.5%). In Hungary, in such regions, the change in the college ratio is 4.6% (s.e. 4.7%) while the change in the skill premium is 8.0% (s.e. 3.5%). Therefore, in low firm density labor markets, the implied firm-specific labor

[34] According to the meta-analysis by Sokolova & Sorensen (2018), the median firm-level labor supply elasticity is around 1.7. Recent quasi experimental studies (e.g. Caldwell & Oehlsen 2018, Cho 2018, Kroft et al. 2020, Dube et al. 2017) find estimates between 2 and 5 (see more details in Bassier et al. 2020).

supply elasticity is around 0.1 in Norway and 0.6 in Hungary, which suggests that firms face quite inelastic labor supply and so they have substantial wage-setting power in these regions. In contrast, in the most dense areas, the skill ratio increases by 8.4% (s.e. 2.4%), the skill premium by 2.3% (s.e. 1.4%) in Norway, while the skill ratio increases by 22.9% (s.e. 11.5%) and the skill premium by 5.1% (s.e. 3.9%) in Hungary. The implied firm-specific labor supply elasticity is 3.6 for Norway and 4.4 for Hungary. This suggest that wage setting power is more limited at high firm density areas. Overall, these findings corroborate a key prediction of our theoretical model: the relative changes in the skill ratio and the skill premium are related to our proxy of firm-specific labor supply elasticities. Furthermore, these geographic differences imply that rural and urban labor markets can be affected quite differently by technological change.

So far we have assumed that firms are atomistic and so they do not consider the impact of their actions on other firms' behavior. Not taking into account such strategic interactions might be unrealistic for larger firms or whenever only a few firms operate in a labor market. Even worse, innovation itself might affect the market power of firms, which could explain the change in the skill premium and the skill ratio even in the absence of changes in skill-biasedness (Berger et al. 2019a). In Appendix D we study whether we see any impact of innovation on subsequent college market share, non-college market share, and relative markdown (see Equation (8)). We find no indication that these proxies of market power change following innovation, except when we use a very narrow definition of the market.[35] We also apply the model of Deb et al. (2020) to calculate the impact of changes of market power following innovation on the skill premium, and find that this impact at most very limited (see the details in Appendix D).

5.3 The Effects of an R&D Tax Credit Policy on Skill Demand

So far we have studied the change in skill premium and skill ratio following innovation. To complement our earlier results, in this section we provide further evidence by exploiting the introduction of an R&D tax credit—a policy-induced change in the cost of a crucial innovation input—in Norway.

In 2002 the government introduced a tax credit that lowered the marginal cost of investing in R&D for a subset of firms. In particular, firms were allowed to deduct up to 20% of their R&D expenses up to a threshold of NOK 4 million (approx 450,000 USD). This implied a reduction in the marginal cost of R&D investments for firms investing less than that threshold. We use a difference-in-differences strategy to compare firms whose marginal cost was affected by the policy to a control group with unaffected firms. This empirical design follows closely that of Bøler et al. (2015) and Bøler (2015). We classify a firm as treated if its average annual R&D expenditure is below 4 million NOK in the pre-tax credit years 1998-2001. We compare these firms to those investing between 4 and 12 million NOK in R&D prior to the policy change. We also restrict the sample to firms with at least 50 employees, as small firms rarely invest in R&D. More details, and sensitivity checks regarding the threshold for the control group, are presented in Appendix A.7.[36]

[35] The narrow definition of the market is at the 3-digit skill-industry-district level. This is a narrower market definition than that of Berger et al. (2019a) who use US industry-commuting zones. Note that our districts are substantially smaller local units than the average US commuting zone.

[36] Note that the first CIS survey was conducted in 2004, so we use another data source, the R&D survey, which goes

Panel A of Figure 3 shows the growth in total log R&D investments relative to the pre-reform year 2001 for treated (solid line) and control firms (dashed line). Treated and control firms follow parallel trends before the reform. However, this trend breaks exactly in 2002, when the tax credit was introduced. The policy led to a 50-100% increase in R&D expenditure among treated firms. Panels B and C show the evolution of the college employment share and the college-to-non-college wage ratio, respectively. The graphs highlight that the increase in R&D expenditures was accompanied by a medium-term increase in the college employment share and in the (raw) college skill-premium among treated relative to control firms. In line with our earlier findings about shorter term effects of innovation (Table 3, columns (6) and (7)), these patterns suggest that it takes time to translate the increased R&D expenditure into actual changes in technology.

Next, we employ a difference-in-differences strategy to estimate the effects of the R&D tax credit on the college employment share and the college premium in treated firms. We run the following regression to assess the impact on the college premium:

$$\ln y_{jt} = \delta Treat_j \times Post_t + \theta_j + \varsigma_{kt} + \epsilon_{jt}, \tag{11}$$

where y_{jt} are various firm-level outcomes (e.g. college share) of firm j at time t, $Treat_j$ is an indicator variable taking the value one if the firm is defined as treated according to the definition above, $Post_t$ is an indicator variable taking the value one for the years following the introduction of the tax credit in 2002, and ς_{kt} reflects industry-year fixed effects. We estimate the regression equation using data for the years 1998-2012, but leaving out the two years immediately following the introduction of the policy.[37]

Columns (1)-(3) in Table 7 show the estimated δ from Equation (11). We find that following the introduction of the R&D tax credit, treated firms increased their college employment share by 8.9 (s.e. 3.1) percent, and the college to non-college employment ratio by around 10.4 (s.e. 4.7) percent, compared to control firms.[38] These findings corroborate the findings of Bøler (2015) who, similarly to us, documented an increase in skill ratio following the introduction of this R&D tax credit policy.

Nevertheless, as we have discussed above, the increase in the skill ratio does not necessarily imply that technological change is skill-biased. Therefore, we also estimate the change in skill premium using a modified version of Equation (9):

$$\ln wage_{ijt} = \delta^u Treat_j \times Post_t + \delta^s Treat_j \times Post_t \times College_i + \gamma X_{ijt} + \eta_i + \varphi_j + \varsigma_{kt} + \varepsilon_{ijt} \tag{12}$$

where $\ln wage_{ijt}$ is the wage for individual i at firm j at time t, X_{ijt} are Mincer-variables, η_i are person effects, φ_j are firm fixed effects, while ς_{kt} are skill-specific time effects (and so they absorb $College_i$, $Post_t$ and $Post_t \times College_i$).

back to periods before the policy reform.

[37] These two years are omitted since it likely takes some time to turn the increase in R&D, an input of the innovation process, into an increase in innovation output, the actual technological change.

[38] Note that the dependent variables are in logs in these regressions and so the coefficients already reflect percent changes. The estimates in Table 5 are in levels and so they reflect percentage point changes. Nevertheless, once we express those in percent changes we find that the college share increased by 5.5 percent and the college-to-non-college ratio by 5.7 percent.

In columns (4)-(5) of Table 7 we report δ^s. Column (4) shows that following the introduction of the tax credit, the college wage premium increased by 5.9 (s.e. 2.8) percent in treated relative to control firms. Column (5) reports estimates when we include worker fixed effects in the regression and so we control for the change in the composition of the workforce even in terms of unobservables. The point estimate is similar to our benchmark estimates on the effect of innovation (3.1% here vs. 4.5% in Table 2) though it is more noisily estimated here. The estimate is not statistically significant from zero, but we can rule out a large (e.g. more than 3%) drop in skill-premium.

To sum up, we find a clear increase in the skill ratio in response to the tax credit driven increase in R&D spending. At the same time, we can also rule out a significant fall in the skill premium even after we control for the change in workers' composition. These findings together indicate that the 2002 tax credit led to technological changes that favored skilled workers. Later we will use these estimates to quantify the contribution of the tax policy to aggregate inequality (see Section 6.2). It is also worth emphasizing that by documenting the changes in wages and employment in response to R&D-driven technological change in general we complement existing evidence on the impact of specific technologies (such as broadband internet) (see Akerman et al. 2015).

5.4 Heterogeneity

In this section we compare different forms of technological change in terms of the skill bias involved. This will help us understand whether only innovations involving R&D or high novelty value are skill biased, or whether firms' skill demand changes even after technology adoption. This question is also linked to the debate about the skill bias of organizational changes as opposed to technical changes (Caroli & Van Reenen 2001). To do so, we investigate the heterogeneity both in terms of the skill premium (Table 6) and the skill ratio (Appendix Section A.6), and propose a way to back out the skill bias based on these two sets of estimates.

Here we describe the skill premium results in more detail. The first column of Table 6 reports our benchmark estimates from column (4) of Table 2. These estimates include worker fixed effects in Norway and were run on the matched sample in Hungary. Column (2) investigates whether innovation by firms conducting R&D is more skill biased than non R&D-based innovation. We study this by including both the basic innovation variable—capturing the effect of non-R&D innovation—and its interaction with an indicator variable showing whether the firm conducts R&D—capturing the additional effect of R&D innovation. The regressions suggest that non-R&D innovation is skill biased in both countries: it has a significant positive effect on the skill premium (and also on skill share, see Appendix Table A.8). Second, R&D innovation seems to be more skill biased than non-R&D innovation: all the coefficients are positive and R&D innovation leads to a significantly higher increase in the wage premium. The relative difference between R&D and non-R&D innovation is much larger in Norway compared to Hungary. Column (3) investigates whether the novelty value of innovation matters. We capture novelty value of innovation with an indicator variable measuring whether it is new to the firm's market. The coefficients of this variable are small and insignificant, suggesting that 'new to the market' innovation is similarly skill biased to other innovations. These results suggest that even low-novelty, non-R&D driven firm-level innovation activities are skill biased and they contribute to the increase in college premium.

Column (4) compares innovation that directly involves technical aspects (product and process) with organizational changes. Note that a firm can conduct both at the same time; therefore, we introduce separate dummies for these two types of innovation. We find that both lead to an increase in the skill premium (and also in the skilled share, see Appendix Table A.8). This reinforces the conclusions of Caroli & Van Reenen (2001) regarding the importance of organizational changes in skill-biased technological change. The magnitude of the change in skill premium after the two types of innovation are similar in Norway, while in Hungary innovations with technical aspects lead to a considerably larger increase in the skill premium. This suggests that organizational changes play a less prominent role in less advanced economies where technological adoption drives innovation activities.

In column (5) we further distinguish between product and process innovation within innovation activities with technical aspects (see footnote 19 for the exact definitions). The point estimate is higher for product innovation in all specifications, even if not significantly, providing some evidence that product innovation is more skill biased than process innovation.

Finally, in column (6) we study whether the skill bias of innovation depends on the technology type of the sector. We classify industries into four groups: high- and low technology manufacturing, and high and low knowledge intensive services.[39] In Norway, the point estimates are very similar in the four sectors, showing that innovation is skill-biased both in manufacturing and services and its effect is largely independent from the technology level of the industry. In Hungary the coefficients are very noisy, but interestingly there seems to be a sharp contrast between manufacturing and services, with no evidence for skill biased technological change in the latter.

Combining the results on the change in skill premium and in skill ratio allows us to back out the average effect of innovation on firm-level skill bias, $\ln \frac{\theta}{1-\theta}$, for the different innovation types. By differentiating, rearranging, and averaging Equation (5), we get:

$$\Delta \ln \frac{\theta}{1-\theta} \equiv \overline{\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}} = \overline{\Delta \ln \frac{w_{H_{jt}}}{w_{L_{jt}}}} + \frac{1}{\sigma} \overline{\Delta \ln \frac{H_{jt}}{L_{jt}}} \tag{13}$$

where $\overline{\Delta \ln \frac{w_{H_{jt}}}{w_{L_{jt}}}}$ is the average change in skill premium following firm-level technological change, and $\overline{\Delta \ln \frac{H_{jt}}{L_{jt}}}$ is the average change in skill ratio following technological change. For the elasticity of substitution between high- and low-skilled labor, σ, we use estimates from the literature, such as $\sigma = 2.94$ from Acemoglu & Autor (2011b).[40]

[39] We use the Eurostat's categorization for this exercise. Manufacturing industries are classified based on the R&D intensities of industries. We consider Eurostat's "High-tech" and "Medium High-tech" industries to be High-tech. These are NACE rev 2 categories 21, 26, 30.3, 20, 25.4, 27, 28, 29, 30 (exl. 30.1 and 30.3) and 32.5. We consider all other manufacturing as low-tech. Knowledge intensive high-tech services are defined based on the share of college educated workers, and the relevant NACE rev 2 codes are: 59-63 and 70. We consider all other non-manufacturing industries sampled by the CIS as not knowledge intensive services.

[40] The estimates are not sensitive to the particular value of σ as long as they are within the range of existing estimates in the literature (i.e. 1 to 10). Furthermore, the $\sigma = 2.94$ is what is implied by the aggregate changes in skill premium, in skill ratio and in skill demand calculated based on that $\sigma = 2.94$ (see the details in Section 6.1). So that level of σ gives us an internally consistent estimate on skill demand.

The results of this procedure are reported in Figure 4. The results largely confirm our earlier findings: non-R&D and low-novelty innovations are also skill biased and both organizational changes and product and process innovation are skill biased. Another characteristic feature is that the difference between R&D and non-R&D innovation is much smaller in Hungary, showing that R&D does not necessarily generate a larger skill-biased change than non-R&D innovation in countries farther from the technology frontier.

6 Quantifying the Effects of Firm-level Technological Change on Inequality

So far we have studied the change in firm-level skill premium and skill ratio following an innovation activity or technological change. In this section we quantify the contribution of firm-level technological change to the aggregate increase in college premium. To do this we impose the following structure on wages:

$$\ln w_{it} = \alpha_t + \psi_i + \ln w_{Sj(i,t)} + \varepsilon_{it} \tag{14}$$

where i denotes workers and j denotes firms, ε_{it} is a mean zero error term. The ψ_i captures workers' skills that are portable across firms and are not affected by firm-level technological change (at least in the short term). The term $\ln w_{Sj(i,t)}$ represents the skill-group (S) specific firm-level wage premium that firm j pays. Let us define the aggregate or economy-wide college premium as the difference between the mean wages, $\overline{\ln w_{H_t}} = \alpha_t + \frac{1}{H_t}\sum_{i\in H}\psi_i + \frac{1}{H_t}\sum_{i\in H}\ln w_{Hj(i,t)}$ and $\overline{\ln w_{L_t}} = \alpha_t + \frac{1}{L_t}\sum_{i\in L}\psi_i + \frac{1}{L_t}\sum_{i\in L}\ln w_{Lj(i,t)}$.

We derive in Appendix F that the contribution of firm-level technological change to the change in the aggregate college premium has the following two terms:

$$\begin{aligned}\Delta\Theta \equiv \Delta\left(\overline{\ln w_{H_t}} - \overline{\ln w_{L_t}}\right) \quad = \quad & \underbrace{\sum_j \left(\frac{H_{jt+1}}{H_{t+1}} - \frac{H_{jt}}{H_t}\right)\ln w_{Hjt+1} - \sum_j \left(\frac{L_{jt+1}}{L_{t+1}} - \frac{L_{jt}}{L_t}\right)\ln w_{Ljt+1} +}_{\text{Reallocation effect}} \\ & + \underbrace{\sum_j \frac{H_{jt}}{H_t}\left(\ln w_{Hjt+1} - \ln w_{Hjt}\right) - \sum_j \frac{L_{jt}}{L_t}\left(\ln w_{Ljt+1} - \ln w_{Ljt}\right)}_{\text{Wage premium effect}}.\end{aligned} \tag{15}$$

where $\Delta\Theta$ denotes the aggregate change in skill bias. The first term captures the reallocation of workers between firms paying different wages. As our empirical analysis demonstrated, firms introducing new technologies hire more skilled workers, which leads to a reallocation of high skilled workers to innovative firms. Furthermore, innovative firms pay higher wages and so reallocation of skilled workers to these firms increases the economy-wide college premium.

The second term captures the change in the skill premium within innovative firms. As our empirical analysis has demonstrated, firms adopting new technologies increase the wage premium of their college workers, which, in itself contributes to wage inequality. In Appendix F, we derive the exact formula for the change in wage premium and show that it can be approximated with

$$\text{Wage premium eff.} \quad \approx \quad \underbrace{\vartheta_{Hjt}^{inn} \times \Delta \ln \frac{\theta}{1-\theta}}_{\text{Direct effect of skill bias}} \tag{16}$$

where $\vartheta_{Hjt}^{inn} \equiv \sum_{j \in inn} \frac{H_{jt}}{H_{t+1}}$ is the share of skilled worker at innovative firms and $\Delta \ln \frac{\theta}{1-\theta}$ is the change in skill bias at innovative firms calculated based on Equation (13) in previous Section.[41]

We explain our empirical implementation in detail in Appendix Section F.1. The results from this exercise are presented in Table 8, which shows the contribution of firm-level technological change to inequality over a 10-year period. Column (1) shows the contribution of all types of technological change—the contribution of firms conducting any type of innovation—to inequality for both countries. The first row is the reallocation effect, which contributed to the increase in skill premium by 0.52 and 3.74 pp. during a 10-year period in Norway and Hungary, respectively. The wage premium effect was 5.58 pp in Norway and 10.09 pp in Hungary.

The total effect is the sum of the reallocation and wage premium effects. Our estimates imply that technological change contributed by 6.1 and 13.83 percentage points to the increase in the economy-wide skill premium over a 10-year period in Norway and Hungary, respectively. The magnitude of this effect is not sensitive to the specific value of σ used for this exercise.[42] The bulk of the contribution comes from the wage premium effect, suggesting that innovation contributes to the economy-wide skill premium via increased wages in innovative firms rather than the reallocation of workers to those firms. The higher contribution in Hungary suggests that technological change farther from the frontier generates more skill bias than the technological change closer to the technological frontier. This finding is also corroborated in a simple cross-country analysis presented in Section 6.1.

We also study the role that different forms of technological changes play in Table 8 and Figure 5. Let us start with columns (2) and (3) of Table 8 and row (1) of Figure 5, which consider R&D and non-R&D innovation. There is a characteristic difference between the two countries: while R&D conducting firms are responsible for 85% of the total increase in inequality in Norway, this number is only 46% in Hungary. This difference has two sources. First, R&D innovation is considerably more skill biased than non R&D-based innovation in Norway, while the difference between the two types of innovation is small in Hungary. Second, R&D firms have a higher market share in Norway.

Columns (4) and (5) in Table 8 and row (2) of Figure 5 compare new-to-market and low-novelty innovation. In Norway, 75% of the aggregate contribution comes from new-to-market innovation, while this number is only 28% in Hungary. The difference is mainly explained by the small prevalence of new-to-market innovation in Hungary compared to Norway.

[41] As we discuss in Appendix Section F.1, the wage premium effect will contain two other terms. Nevertheless, it turns out that those terms will be very small empirically. In our calculations presented in Table 8, we take into account those terms as well but that has a negligible effect on our estimates (see Appendix Table F.3).

[42] In Norway, the contribution per year changes from the baseline 6.1 to 7.5 and 5 percentage points when using $\sigma = 1.6$ and $\sigma = 10$, respectively. In Hungary, the annual values change to 15.7 and 9.6 percentage points when $\sigma = 1.6$ and $\sigma = 10$, respectively (see Appendix Table F.4).

Finally, columns (6)-(8) in Table 8 and row (3) of Figure 5 analyze firms conducting innovation with only technical aspect, only organizational change or combining the two. Firms conducting both types of innovation generate the bulk of the contribution in both countries, both because of the higher skill bias of this type of innovation and also because of the large market share of firms conducting both types.

These findings underline the higher importance of technology adoption—either captured by non-R&D or low-novelty innovation—in Hungary than in Norway, where the economy-wide skill premium is mainly driven by R&D-based, higher novelty innovation. Furthermore, firms conducting both technical and organizational innovation contribute most to the increase in inequality, suggesting strong complementarity between the two types of innovation.

6.1 Economy-wide Skill Premium, Skill Ratio and Skill-Bias

How is the estimated contribution of technological change to economy-wide college premium related to the actual changes observed in the data? As we described in the Introduction, the college premium has been falling both in Norway and in Hungary. In particular, the actual skill premium in the period that we studied here declined from 31 to 20% in Norway, and from 110 to 95% in Hungary.[43] These trends seem to contradict our estimates that predict an increase in inequality in this period.

Nevertheless, the fall in college premium coincided with a significant increase in the college to non-college ratio, which has been increasing from 0.49 to 0.75 in Norway, and from 0.16 to 0.32 in Hungary. In fact, the relative increase in skilled workforce can itself explain the fall in the college premium if the the elasticity of substitution between the two skill groups, σ, is large enough. In Table 9, we calculate the σ that is needed to reconcile the change in the college premium and ratio without any skill bias in technological change for Norway (Panel A) and Hungary (Panel B). In particular, we use the following equation to back out σ (and assume, for now, that $\Delta\Theta = 0$):

$$\Delta \ln \frac{w_{H_t}}{w_{L_t}} = \Delta\Theta + \frac{1}{\sigma}\Delta \ln \frac{H_t}{L_t} \tag{17}$$

The implied σ without technological change is 4.9 for Norway and 9.4 for Hungary. We also do the same exercise for the United States (Panel C), which also experienced a fall in college premium and an increase in college ratio in this period.[44] The implied elasticity without technological change is 11.75 for the United States.

In all three countries the σ that is needed to reconcile the changes in college premium and college ratio in absence of skill bias in technological change is considerably larger than the elasticity implied by earlier periods (see Acemoglu & Autor 2011a). However, once we substitute our estimated contribution of technological change to the change in college premium (the total effect, $\Delta\Theta$, from Table 8)[45], we get

[43] These data come from the OECD Education at a Glance 2014 and 2020. Since the college premium is missing for 2000, we study the period between 2005 and 2015 for Norway, and the 2000-2015 period for other countries.

[44] The college premium decreased from 76 to 71%, and the college ratio increased from 0.57 to 0.81 between 2000 and 2015.

[45] The estimates in Table 8 are for a 10-year period, so we multiply those changes by 1.5 in Hungary and in the

that the implied elasticities of substitution are 2.47 in Hungary and 2.87 in Norway (see the second rows of Table 9), which are very close to the elasticity of substitution found in Acemoglu & Autor (2011a). Furthermore, if we apply the estimated skill bias contribution from Norway—which is more similar to the USA in terms of its distance from the technology frontier—to the USA, we find that the implied elasticity is 2.80, which is again very much consistent with the long-term evidence from the United States. To sum up, the estimated technological change seems to be consistent with the observed evolution of the economy-wide college premium.

Another interesting finding is the difference in the contributions of technological change to inequality between Norway and Hungary. We found that in Hungary, which is farther from the technology frontier and mainly adopting technologies used in more developed countries, the skill bias is larger than in Norway, which is closer to the technology frontier. Is this simply a coincidence? By observing changes in the skill premium and skill share in a country and assuming a specific value for σ, we can back up the implied contribution of skill-biased technological change using equation (17). We classify countries into groups according to their European Innovation Scoreboard, which measures research and innovation performance of European countries. Assuming $\sigma = 2.96$, we find that the implied contribution of skill bias is 8.5% for innovation leaders, 13.9% for strong innovators (the group which includes Norway) and 21.3% for moderate innovators (the group that includes Hungary).[46] This shows that the pattern of the contribution being larger in Hungary compared to Norway is not atypical, and reflects the substantial skill bias involved in technology adoption.

6.2 The Effect of the R&D Tax Credit on the Economy-wide College Premium

We can also apply our approach to quantify the contribution of the R&D tax credit policy, described in Subsection 5.3, to the college premium (see the details in Appendix Section F.2). We estimate that the R&D tax credit reform increased the economy-wide college premium by 1.39 percentage points in the long-term. This highlights that policies encouraging innovation can have substantial effects on inequality.

7 Conclusion

This paper documents that innovation activities and technological change are associated with an increase in skill demand in Norway and Hungary. Our approach directly infers skill bias from firm-level technological change. We exploit an exceptionally rich survey data, the CIS, which provides self-reported measures of firm-level technological change. We identify and quantify the extent to which firm-level technological change is skill biased by estimating the change in both the skill ratio

United States to translate them into a 15-year period change.

[46] Innovation leaders: Finland, Denmark, Sweden, Switzerland, Israel; Strong Innovators: Norway, Austria, Germany, UK, Belgium, France, Portugal, Ireland; Moderate Innovators: Hungary, Italy, Czechia, Spain and Turkey. The fourth category in the Innovation Scorecard is "Modest innovator", but there were no countries in this group with OECD data available.

and skill demand following innovation. We find that innovation is a key force behind the recent trends of inequality. The large contribution of skill-biased technological change might be surprising given the considerable fall in college premium observed recently in many countries. Nevertheless, we demonstrate that the fall in the college premium simply reflects that in the most recent periods the race between education and technology (Goldin & Katz 2010) was won by education. Our estimates imply that technological change still plays a prominent role in the evolution of the college premium.

Comparing the two countries, interestingly, we find that the skill demand increase was substantially larger in Hungary, the country farther away from the technological frontier. Our findings demonstrate that technology adoption can be a very important source of rising inequality in countries farther from the technological frontier. These results highlight that the nature of technological progress matters for shaping inequality.

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

Acemoglu, D. & Autor, D. (2011b), Skills, tasks and technologies: Implications for employment and earnings, in 'Handbook of labor economics', Vol. 4, Elsevier, pp. 1043–1171.

Acemoglu, D., Lelarge, C. & Restrepo, P. (2020), Competing with robots: Firm-level evidence from france, in 'AEA Papers and Proceedings', Vol. 110, pp. 383–88.

Acemoglu, D. & Restrepo, P. (2020), Unpacking skill bias: Automation and new tasks, in 'AEA Papers and Proceedings', Vol. 110, pp. 356–61.

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

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

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.

Bassier, I., Dube, A. & Naidu, S. (2020), Monopsony in movers: The elasticity of labor supply to firm wage policies, Technical report, National Bureau of Economic Research.

Beaudry, P., Doms, M. & Lewis, E. (2010), 'Should the personal computer be considered a technological revolution? evidence from us metropolitan areas', Journal of Political Economy **118**(5), 988–1036.

Berger, D., Herkenhoff, K. & Mongey, S. (2019a), 'Labor market power', IZA Discussion Paper .

Berger, D. W., Herkenhoff, K. F. & Mongey, S. (2019b), Labor market power, Technical report, National Bureau of Economic Research.

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 .

Bøler, E. A., Moxnes, A. & Ulltveit-Moe, K. H. (2015), 'R&d, international sourcing, and the joint impact on firm performance', American Economic Review **105**(12), 3704–39.

Bound, J. & Jonson, G. (1992), 'Changes in the structure of wages in the 1980's: An evaluation of alternative explanations', The American Economic Review **82**(3), 371–392.

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.

Caldwell, S. & Oehlsen, E. (2018), 'Monopsony and the gender wage gap: Experimental evidence from the gig economy', Massachusetts Institute of Technology Working Paper .

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.

Cho, D. (2018), 'The labor market effects of demand shocks: Firm-level evidence from the recovery act'.

Ciccone, A. & Hall, R. E. (1996), 'Productivity and the density of economic activity', The American economic review pp. 54–70.

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.

De Loecker, J., Eeckhout, J. & Unger, G. (2020), 'The rise of market power and the macroeconomic implications', The Quarterly Journal of Economics **135**(2), 561–644.

Deb, S., Eeckhout, J., Patel1, A. & Warren, L. (2020), 'The contribution of market power to wage inequality', mimeo .

DiNardo, J. E. & Pischke, J.-S. (1997), 'The returns to computer use revisited: Have pencils changed the wage structure too?', The Quarterly Journal of Economics **112**(1), 291–303.

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.

Doms, M., Dunne, T. & Troske, K. R. (1997), 'Workers, wages, and technology', The Quarterly Journal of Economics **112**(1), 253–290.

Dube, A., Manning, A. & Naidu, S. (2017), 'Monopsony and employer mis-optimization account for round number bunching in the wage distribution', Unpublished manuscript .

Dunne, T., Foster, L., Haltiwanger, J. & Troske, K. R. (2004), 'Wage and productivity dispersion in united states manufacturing: The role of computer investment', Journal of Labor Economics **22**(2), 397–429.

Evangelista, R. & Vezzani, A. (2010), ‘The economic impact of technological and organizational innovations. a firm-level analysis’, Research Policy **39**(10), 1253–1263.

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

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E. et al. (2019), ‘Toward understanding the impact of artificial intelligence on labor’, Proceedings of the National Academy of Sciences **116**(14), 6531–6539.

Garin, A. & Silvério, F. (2018), ‘How responsive are wages to demand within the firm? evidence from idiosyncratic export demand shocks’, Mimeo .

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

Graetz, G. & Michaels, G. (2018), ‘Robots at work’, Review of Economics and Statistics **100**(5), 753–768.

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.

Griliches, Z. (1957), ‘Hybrid corn: An exploration in the economics of technological change’, Econometrica, Journal of the Econometric Society pp. 501–522.

Haanwinckel, D. (2018), ‘Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution’, Unpublished manuscript .

Hjort, J. & Poulsen, J. (2019), ‘The arrival of fast internet and employment in africa’, American Economic Review **109**(3), 1032–79.

IMF (2015), ‘Norway: selected issues’, IMF Staff Country Reports (2015/250).

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.

Kline, P., Petkova, N., Williams, H. & Zidar, O. (2019), ‘Who Profits from Patents? Rent-Sharing at Innovative Firms*’, The Quarterly Journal of Economics **134**(3), 1343–1404.

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

Kroft, K., Luo, Y., Mogstad, M. & Setzler, B. (2020), Imperfect competition and rents in labor and product markets: The case of the construction industry, Technical report, National Bureau of Economic Research.

Krueger, A. B. (1993), ‘How computers have changed the wage structure: evidence from microdata, 1984–1989’, The Quarterly Journal of Economics **108**(1), 33–60.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V. & Violante, G. L. (2000), ‘Capital-skill complementarity and inequality: A macroeconomic analysis’, Econometrica **68**(5), 1029–1053.

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'.

Lopez-Rodriguez, J. & Martinez-Lopez, D. (2017), 'Looking beyond the r&d effects on innovation: The contribution of non-r&d activities to total factor productivity growth in the eu', Structural Change and Economic Dynamics **40**, 37–45.

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.

Manning, A. (2013), Monopsony in motion: Imperfect competition in labor markets, Princeton University Press.

McFadden, D. et al. (1977), Quantitative methods for analyzing travel behavior of individuals: some recent developments, Vol. 474, Institute of Transportation Studies, University of California Berkeley, CA.

Milgrom, P. & Roberts, J. (1990), 'The economics of modern manufacturing: Technology, strategy, and organization', The American Economic Review pp. 511–528.

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

Nelson, R. R. & Phelps, E. S. (1966), 'Investment in humans, technological diffusion, and economic growth', The American Economic Review **56**(1/2), 69–75.

Nergaard, K. (2014), 'Trade unions in norway'.

Rigó, M. (2012), 'Estimating union-non-union wage differential in hungary', Unpublished PhD dissertation chapter. Central European University .

Saez, E., Schoefer, B. & Seim, D. (2019), 'Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in sweden', American Economic Review **109**(5), 1717–63.

Sokolova, A. & Sorensen, T. (2018), 'Monopsony in labor markets: A meta-analysis. iza discussion papers 11966', Institute for the Study of Labor (IZA) .

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

Stansbury, A. & Summers, L. H. (2020), The declining worker power hypothesis: An explanation for the recent evolution of the american economy, Technical report, National Bureau of Economic Research.

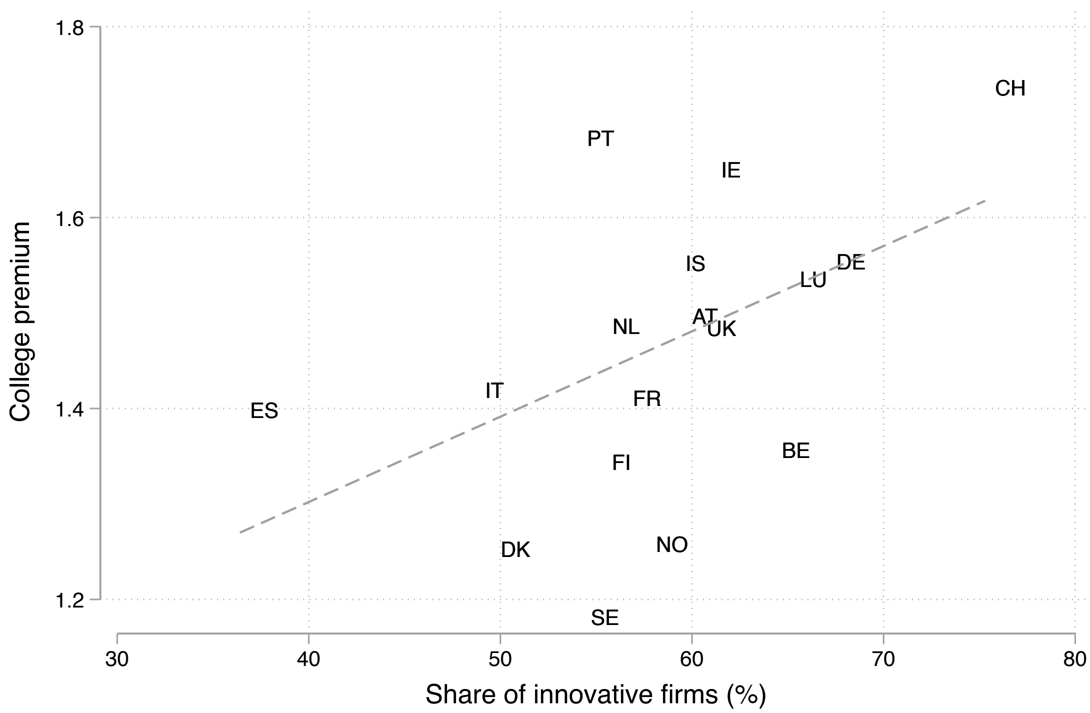
Tonin, M. et al. (2009), 'Employment protection legislation in central and east european countries', SEER-South-East Europe Review for Labour and Social Affairs (04), 477–491.

Van Reenen, J. (1996), 'The creation and capture of rents: wages and innovation in a panel of uk companies', The Quarterly Journal of Economics **111**(1), 195–226.

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

Figures

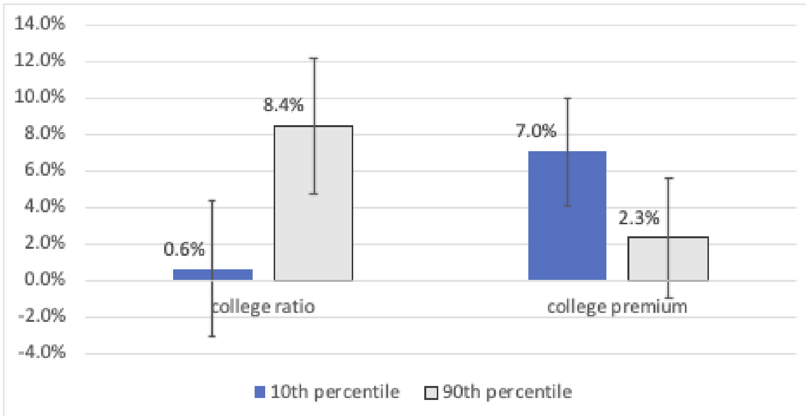
Figure 1: Share of Innovative Firms and the College Premium: Cross-Country Evidence



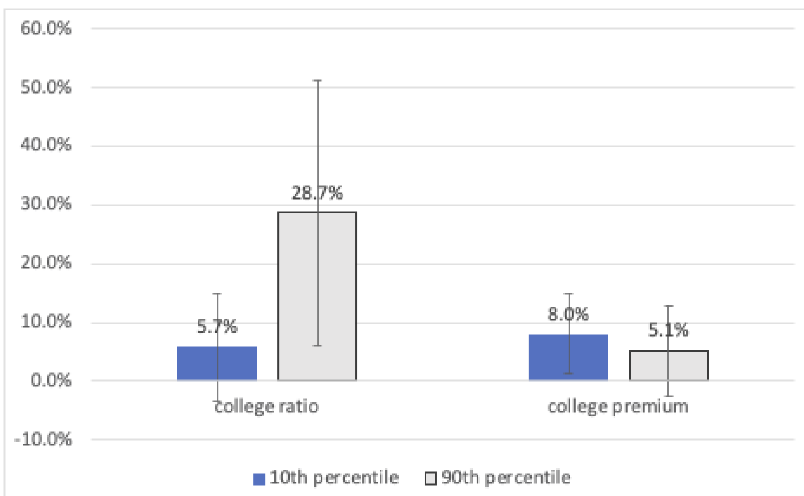
Notes: This figure shows the cross-country relationship between the college premium and the share of innovative firms in 2014. Innovative firms are those firms changing their technology between 2012 and 2014 by introducing any new or significantly modified product/service/process/organizational change. The data comes from Eurostat. The innovation variable is from the 2014 Community Innovation Survey, while the college premium comes from the Structure of Earnings Survey.

Figure 2: Change in Skill Demand Following Technological Change by Firm Density

Panel A: Norway



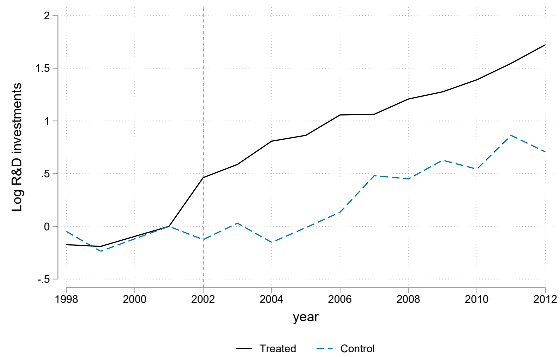
Panel B: Hungary



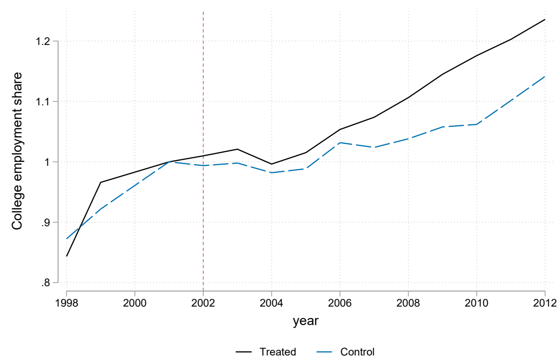
Notes: This figure shows percent changes in the college to non-college ratio and in the college wage premium following firm-level technological change for local areas with low (10th percentile) and high (90th percentile) firm density. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. Firm density is measured as the log number of firms per square kilometer. This variable proxies the dispersion of workers' idiosyncratic preferences for working at a particular firm, which is tightly linked to the firm-specific elasticity of labor supply in the model. We obtain the percent change in the skill ratio values by adding an interaction term between the innovation variable and log firm density in the local area to our benchmark specification (Table 5 column (3)). The point estimates of the interaction term are reported in Appendix Table A.6, here we report the marginal effect of innovation on the skill ratio at the 10th and 90th percentile of the local area firm density distribution. We transform our estimates from percentage points to percent based on the average H/L value of non-innovative firms in Table 1. We obtain the percent change in college premium by adding an interaction term between the innovation variable and the log firm density in the local area to our benchmark specification (Table 2 column (4)). The point estimates of the interaction term are reported in Appendix Table A.7, here we report the marginal effect of innovation on the skill ratio at the 10th and 90th percentile of the local area firm density distribution. We clustered the standard errors at the firm-level in both regressions. The error bars show the 95% confidence interval around the point estimate.

Figure 3: Change in R&D investments and in Skill Demand Following the Introduction of an R&D Tax Credit Policy in Norway

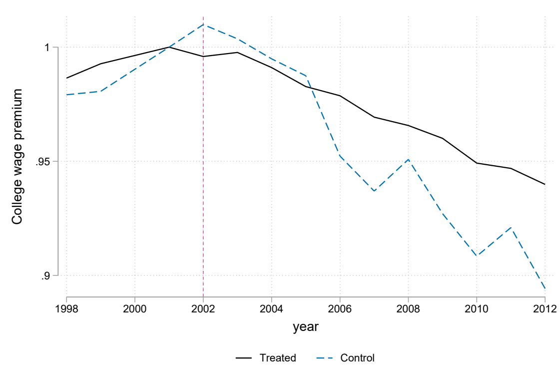
Panel A: Log R&D Investments



Panel B: College Employment Share

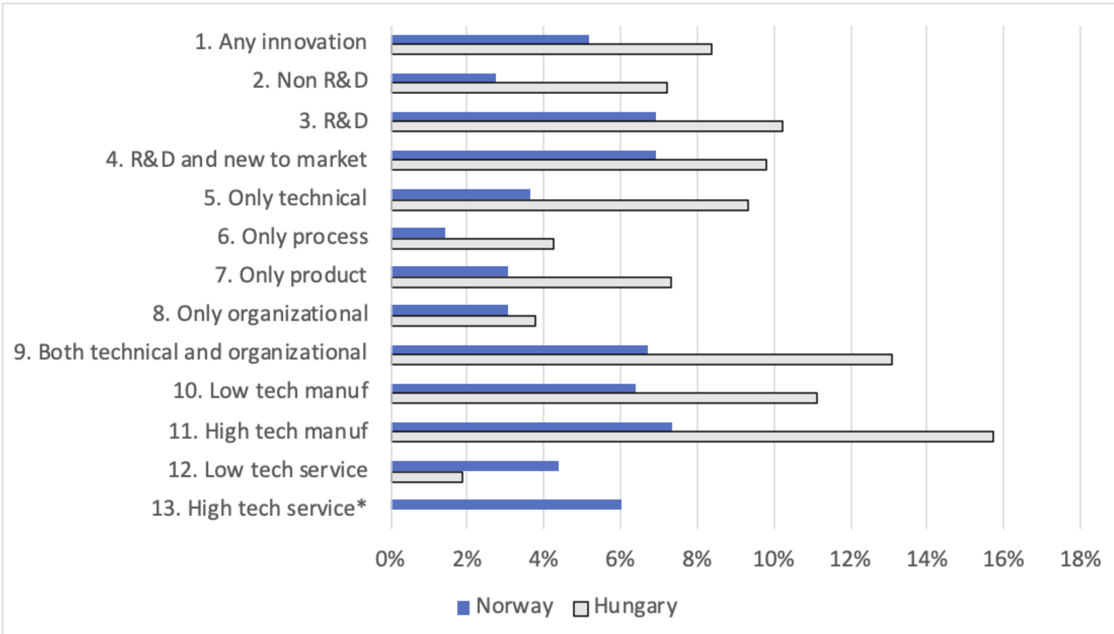


Panel C: College-to-Non-College Wage Ratio



Notes: This figure shows the evolution of R&D investment and the skill ratio following the 2002 introduction of a R&D tax credit policy in Norway. The tax credit allowed firms to deduct up to 20% of their R&D expenses up to a threshold of NOK 4 million (approx 450,000 USD). This implied a reduction in the marginal cost of R&D investments for firms investing less than the threshold. We assign firms to the treated group if they spent less than the policy threshold (NOK 4 million) on R&D prior to the reform. We assign firms to the control group if they reported R&D expenses between NOK 4-12 million prior to the reform. Since R&D investments are mainly relevant for larger firms, we include firms with at least 50 employees. Panels A shows the (log) total R&D investment for the firms in the treated and in the control groups. Panel B shows the average college employment share for the firms in the treated and in the control groups while Panel C shows the average college to non-college wage ratio for the two groups (both weighted by the number of workers). All outcomes are normalized to be 100% (or one) in 2001 (the last year prior to the reform).

Figure 4: Estimated Firm-Level Skill Bias of Different Forms of Technological Change

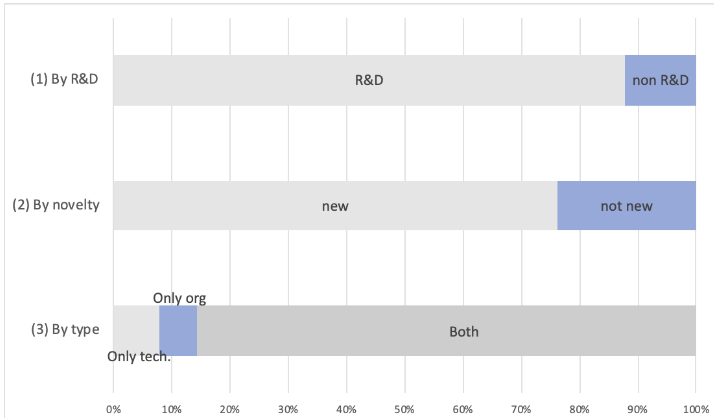


Notes: This figure shows the change in skill bias $\left(\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}\right)$ calculated from the estimated change in skill premium and skill ratio following technological change. The skill premium and skill ratio estimates are from Tables 6 and A.8. We use Equation (13) and $\sigma = 2.96$ (following Acemoglu & Autor 2011a) to obtain the change in skill bias. In particular, we take the estimated (percent) change in the skill premium and add to that $1/\sigma$ times the estimated percent change in the skill ratio. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. Row 1 shows the change in skill bias for firms conducting any type of innovation. We measure different forms of technological change from the detailed questionnaire of the CIS survey on firms' innovation activities. Rows 2 and 3 show the change in skill bias for innovative firms conducting in-house R&D and for other innovators, respectively. Row 4 shows the change in skill bias for firms conducting R&D and introducing novel processes or products that are new to the firms' market, rather than only for the firm. Rows 5, 6, 7 and 8 plot the skill bias for firms conducting only technical (process or product) innovation, only process innovation, only product innovation, or only organizational change, respectively. Row 9 shows the change in skill bias for firms combining technical innovation with organizational change. Rows 10-13 show the change in skill bias for firms operating in various industries. We follow the Eurostat categorization and assign firms to High-tech and Medium High tech manufacturing industries ("High tech manuf."); other manufacturing ("Low tech manuf") ; high-tech knowledge intensive services ("High tech services") and other service industries ("Low tech services"). The blue, filled bars show the change in skill premium for Norway, and the gray bars for Hungary.

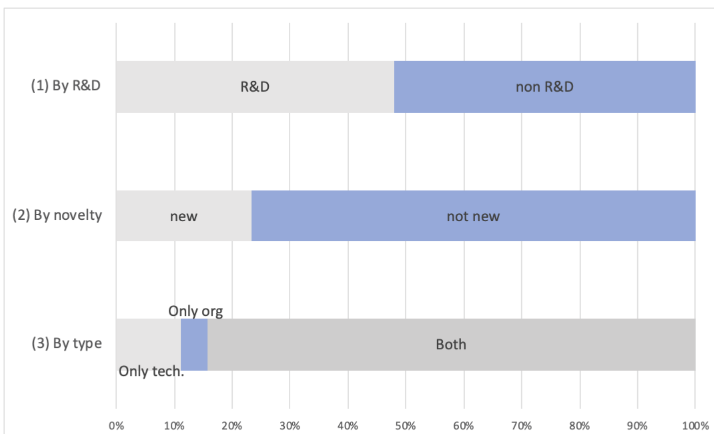
*: there are very few "High tech service firms" in Hungary. As a result, we do not include the outlier (and insignificant) -13.8 percent change in skill bias for High-tech Services in Hungary.

Figure 5: Contribution of Different Forms of Technological Change to the Economy-wide Skill Premium

Panel A: Norway



Panel B: Hungary



Notes: This figure shows the relative contribution of different forms of firm-level technological changes to the economy-wide college premium. Firm-level technological change measured in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. The estimates show the sum of the reallocation effect and the wage premium effect (see Equation (15)) by three, mutually exclusive breakdowns of innovative firms. The first row breaks down the contribution by R&D. We calculate the contribution of R&D conducting innovators and the contribution of innovators not relying on R&D (non R&D). We plot the relative contributions of these two groups of firms. The second row breaks down the contribution by novelty. We calculate the contribution of firms introducing process and/or product innovations that are new to the market (new) and that of other innovators (not new). We plot the relative contributions of these groups. The last row shows the contribution by types of technological change. We calculate the contribution of innovators introducing new products or processes (only technical), of firms conducting only organizational innovation (only organisational), and of firms that combine the two (both). Then plot the relative contributions of these three groups. Further details are provided in Section 6 and Appendix Section F.1. Panel A shows the estimates for Norway and Panel B for Hungary.

Tables

Table 1: Descriptive Statistics: Characteristics of Innovative and Non-innovative firms

	Norway		Hungary	
	Non-innovative	Innovative	Non-innovative	Innovative
Average years of education	12.70	13.41	11.77	12.41
	(1.59)	(1.64)	(1.43)	(1.49)
Share of college graduates	0.19	0.31	0.17	0.28
	(0.25)	(0.28)	(0.22)	(0.24)
College to non-college ratio	0.49	0.87	0.20	0.47
	(0.98)	(1.22)	(0.39)	(0.51)
Log average daily wage rate (EUR)	4.68	4.84	3.07	3.35
	(0.47)	(0.42)	(0.46)	(0.48)
Average age of employees	41.38	41.3	44.04	42.66
	(5.89)	(4.97)	(5.67)	(5.05)
Number of employees	33.50	129.02	146.64	462.30
	(103.24)	(417.63)	(240.17)	(1557.154)
Number of firm-years	16,921	15,528	1,577	971

Notes: This table shows the characteristics of innovative and non-innovative firms in the Community Innovation Survey (CIS). We measure firm-level technological change in the CIS survey. Innovative firms report that they introduced new or significantly modified products/technologies/organization, which are new from their point of view. Non-innovative firms are the rest of the firms in the survey. We report the outcomes in 2012. The table shows the mean of firm-level average years of schooling, the mean of firm-level share of college graduates, the mean of firm-level college to non-college ratio, the mean of firm-level average log daily wage, the mean of firm-level average age of workers, and the mean of firms' number of employees. We report the standard deviation of these variables in parentheses below.

Table 2: Change in the Skill Premium Following Firm-level Technological Change

Panel A: Norway

	(1)	(2)	(3)	(4)	(5)	(6)
Innovation	0.105***	0.090***	-0.024***	-0.011	-0.010	-0.012
	(0.019)	(0.017)	(0.009)	(0.008)	(0.008)	(0.008)
Innovation x College	0.077***	0.068***	0.130***	0.045***	0.035***	0.034***
	(0.025)	(0.025)	(0.016)	(0.010)	(0.011)	(0.011)
Innovation +2 x College					0.014*	0.005
					(0.008)	(0.008)
Innovation +4 x College						0.012
						(0.009)
Mincer variables	No	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes	Yes	Yes
Worker FEs	No	No	No	Yes	Yes	Yes
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373
Firms in CIS	15,530	15,530	15,530	15,530	15,530	15,530
R-squared	0.05	0.07	0.20	0.44	0.44	0.44

Panel B: Hungary

	(1)	(2)	(3)	(4)	(5)	(6)
Innovation	0.201***	0.166***	-0.028**	-0.008	-0.005	-0.006
	(0.022)	(0.019)	(0.012)	(0.009)	(0.010)	(0.011)
Innovation x College	0.085***	0.100***	0.123***	0.067***	0.069***	0.065***
	(0.027)	(0.023)	(0.014)	(0.023)	(0.024)	(0.023)
Innovation +2 x College					0.020	0.008
					(0.026)	(0.023)
Innovation +4 x College						0.014
						(0.021)
Mincer variables	No	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes	Yes	Yes
Matched sample	No	No	No	Yes	Yes	Yes
Observations in CIS	785,443	785,443	785,419	197,065	197,065	197,065
Firms in CIS	6,212	6,212	6,212	1,716	1,716	1,716
R-squared	0.44	0.51	0.71	0.70	0.70	0.70

Notes: This table investigates the change in workers' (log) wages following firm-level technological change. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. We report the estimated coefficients on the innovation dummy, δ^u, and the innovation dummy interacted with whether the individual has a college degree, δ^s, from Equation (9) described in Section 4.1. The "Innovation" dummy indicates whether technological change was introduced according to the current CIS wave or any of the previous two waves. Our primary interest lies in the coefficient of the "Innovation x College" interaction, which shows the extent to which the college premium changes following technological change relative to firms not reporting any technological change. Panel A shows the estimates for Norway, while panel B the estimates for Hungary. All specifications include skill-year fixed effects, representing interactions of primary, secondary, vocational and college dummies with year dummies. Column (1) shows the estimates when including only skill-year (e.g. college-year) fixed effects in the regression. Columns (2)-(6) also include Mincer variables (gender, age, tenure, tenure squared, a dummy for new entrant in both countries and hours worked and a dummy for part-time employees in Hungary where part-time workers are also included in the sample). Columns (3)-(6) add firm fixed effects to the regression. Columns (4)-(6) include worker fixed effects in Norway and apply the matching procedure for Hungary (discussed in detail in Section 4.1 and Appendix Section B.5). Columns (5) and (6) also include pre-trend dummies showing whether the firm innovated in the subsequent CIS wave or the wave after that. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 3: Robustness: Change in Skill Premium Following Firm-Level Technological Change

Panel A: Norway

	(1) Industry and district-year-FE	(2) Industry-district-year FE	(3) Occupation-district-year FE	(4) Industry-occupation-district-year-FE	(5) District-college wage share-year-FE	(6) Short term	(7) Medium term	(8) Firm specific college premium FEs
Innovation	-0.009	-0.010	-0.024***	-0.017**	-0.002	-0.005	-0.012*	-0.018***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.005)
Innovation x College	0.046***	0.046***	0.048***	0.036***	0.022**	0.018**	0.034***	0.035***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.009)	(0.008)	(0.009)	(0.009)
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373
R-squared	0.44	0.44	0.44	0.45	0.46	0.44	0.44	0.46

Panel B: Hungary

	(1) Industry and district-year-FE	(2) Industry-district-year FE	(3) Occupation-district-year FE	(4) Industry-occupation-district-year-FE	(5) District-college wage share-year-FE	(6) Short term	(7) Medium term
Innovation	-0.011	-0.015	-0.006	-0.009	-0.007	-0.010	-0.007
	(0.008)	(0.010)	(0.007)	(0.009)	(0.007)	(0.007)	(0.008)
Innovation x College	0.067***	0.069***	0.063***	0.055***	0.057***	0.059***	0.057**
	(0.024)	(0.026)	(0.016)	(0.019)	(0.018)	(0.022)	(0.022)
Observations in CIS	193,019	192,970	193,797	180,456	194,352	174,102	190,666
R-squared	0.71	0.72	0.78	0.84	0.77	0.70	0.70

Notes: This table shows robustness checks for the results on the change in workers' (log) wages following firm-level technological change presented in Table 2. We report the estimated coefficients on the innovation dummy, δ^u, and the innovation x college interaction, δ^s, from Equation (9). All specifications include skill-year (e.g. college-year) fixed effects, Mincer variables and firm fixed effects. Worker fixed effects are also included in Norway, while we apply the matching procedure for Hungary. Column (1) adds additionally industry-year and district-year fixed effects, Column (2) industry-district-year fixed effects, Column (3) occupation-district-year fixed effects, and Column (4) adds industry-occupation-district-year fixed effects. In column (5) we classify firms into skill ratio quartiles based on their initial year in our sample (the starting year of our analysis or the entry date for firms entering later) and then add quartile-district-year fixed effects to the regression. Column (6) shows short-term changes by defining innovation based on the current CIS wave, while Column (7) shows the medium-term changes by defining innovation using the current CIS and last CIS wave, rather than the previous two waves, as in our main specification. Column (8) includes firm-college fixed effects for Norway by grouping firms into deciles based on their college premium and then we include an additional interaction of firm premium-type deciles with the college dummy in the regression (see the details Section 4.1). Standard errors are clustered at the firm level. Significance levels are: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 4: Change in Skill Premium Following Firm-level Technological Change for Incumbent Workers and for New Entrants

Panel A: Norway

	(1)	(2)	(3)	(4)
Innovation x New entrant	0.093***	0.084***	-0.021**	-0.028***
	(0.016)	(0.015)	(0.011)	(0.008)
Innovation x Incumbent	0.068**	0.075***	-0.023***	0.010
	(0.027)	(0.027)	(0.009)	(0.008)
Innovation x College x New entrant	0.073***	0.061**	0.135***	0.043***
	(0.025)	(0.025)	(0.018)	(0.010)
Innovation x College x Incumbent	0.054*	0.043	0.059***	0.020*
	(0.030)	(0.029)	(0.013)	(0.011)
Skill-year FE	Yes	Yes	Yes	Yes
Mincer variables	No	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes
Worker FEs	No	No	No	Yes
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373
R-squared	0.05	0.07	0.20	0.44

Panel B: Hungary

	(1)	(2)	(3)	(4)
Innovation x New entrant	0.139***	0.135***	-0.026***	-0.016
	(0.024)	(0.020)	(0.008)	(0.010)
Innovation x Incumbent	0.180***	0.159***	-0.005	-0.003
	(0.026)	(0.022)	(0.007)	(0.007)
Innovation x College x New entrant	0.036	0.035	0.086***	0.095***
	(0.029)	(0.025)	(0.016)	(0.025)
Innovation x College x Incumbent	0.049*	0.059**	0.080***	0.043*
	(0.026)	(0.024)	(0.013)	(0.024)
Skill-year FE	Yes	Yes	Yes	Yes
Mincer variables	No	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes
Matched sample	No	No	No	Yes
Observations in CIS	703,539	703,539	703,508	174,102
R-squared	0.461	0.511	0.716	0.704

Notes: This table investigates the change in workers' (log) wages following firm-level technological change for incumbent workers and for new entrants. We start from the benchmark regression Equation (9) and add the innovation dummy interacted with new entrants/incumbent status and the triple interaction terms between innovation x college x new entrants/incumbent status. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. In Norway, incumbents are defined as individuals who had been working at the firm for at least 6 years, and new entrants are all other workers. To make sure that incumbent workers had been present at the firm before innovation started we focus on medium-term effects of innovation (same as in column (7) in Table 3). In Hungary, incumbents are defined as individuals who had been working at the firm for at least 24 months, and new entrants are all other workers. To make sure that incumbent workers had been present at the firm before innovation started we focus on short-term effects of innovation (same as in column (6) in Table 3). Column (1) shows the estimates when including only skill-year (e.g. college-year) fixed effects in the regression. Columns (2)-(4) also include the Mincer variables, columns (3)-(4) add firm fixed effects to the regression and columns (4) include worker fixed effects in Norway and apply the matching procedure for Hungary (discussed in detail in Section 4.1 and Appendix Section B.5). Standard errors are clustered at the firm level. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 5: Change in the Skill Ratio Following Firm-level Technological Change

Panel A: Norway

	(1) College employment share	(2) College to non college employment ratio	(3) log employment
Innovation	0.011***	0.028***	0.053***
	(0.002)	(0.006)	(0.011)
ln capital (d)	0.000	-0.005*	
	(0.001)	(0.003)	
ln value added (d)	-0.006**	-0.003	
	(0.003)	(0.007)	
Industry-year FE	Yes	Yes	Yes
Observations	18,215	17,796	24,945
R-squared	0.06	0.05	0.07

Panel B: Hungary

	(1) College employment share	(2) College to non college employment ratio	(3) log employment
Innovation	0.019**	0.029***	0.030
	(0.008)	(0.008)	(0.020)
ln capital (d)	-0.007	-0.012*	
	(0.007)	(0.007)	
ln value added (d)	-0.007	-0.005	
	(0.008)	(0.009)	
Industry-year FE	Yes	Yes	Yes
Observations	2,153	2,125	2,363
R-squared	0.10	0.16	0.14

Notes: This table shows the relationship between firm-level technological change and subsequent 6-year change in firm-level college employment share (column 1), in college to non-college ratio (column 2), and log employment (column 3). We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. In the table we report the δ coefficients from regression Equation (10). The "Innovation" dummy indicates whether the firm innovated according to the current CIS wave or any of the previous two waves. The other two explanatory variables in columns (1)-(2) are long differences of log capital stock and log value added. In each regression we include the lagged dependent variable preceding the baseline year and industry-year fixed effects. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 6: Change in the College Premium Following Different Forms of Technological Change

Panel A: Norway

	(1)	(2)	(3)	(4)	(5)	(6)
Innov x College	0.045***	0.022*	0.021*			
	(0.010)	(0.012)	(0.012)			
Innov x R&D x College		0.039***	0.036***			
		(0.011)	(0.012)			
Innov x New x College			0.005			
			(0.011)			
Tech. x College				0.033***		
				(0.010)		
Org. x College				0.023**	0.022**	
				(0.009)	(0.009)	
Process x College					0.013	
					(0.010)	
Product x College					0.027***	
					(0.010)	
Innov x Manuf. x College						0.057***
						(0.013)
Innov x HT manuf. x College						0.058***
						(0.018)
Innov x Services x College						0.037***
						(0.012)
Innov x HK services x College						0.057***
						(0.015)
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373	4,804,373
R-squared	0.44	0.44	0.44	0.44	0.44	0.44

Panel B: Hungary

	(1)	(2)	(3)	(4)	(5)	(6)
Innov. x College	0.067***	0.059**	0.059**			
	(0.023)	(0.024)	(0.024)			
Innov. x R&D x College		0.023	0.025			
		(0.028)	(0.028)			
Innov. x New x College			-0.006			
			(0.032)			
Tech. x College				0.073***		
				(0.022)		
Org. x College				0.021	0.009	
				(0.022)	(0.031)	
Process x College					0.031	
					(0.032)	
Product x College					0.060**	
					(0.025)	
Innov. x Manuf. x College						0.071***
						(0.026)
Innov. x HT manuf. x College						0.127***
						(0.038)
Innov. x Services x College						0.027
						(0.035)
Innov. x HK services. x College						-0.114
						(0.089)
Observations in CIS	197,065	197,065	197,065	197,262	197,262	197,262
R-squared	0.70	0.70	0.70	0.70	0.70	0.70

Notes: This table shows the change in workers' (log) wages following different forms of firm-level technological changes. We measure different forms of technological change from the detailed questionnaire of the CIS survey on firms' innovation activities. The table reports regression estimates that extend the benchmark specification (reported in column 4 of Table 2). Column (2) includes a dummy showing whether the innovating firm conducted R&D and column (3) also includes a dummy showing whether the innovation was new for the firms' market rather than only for the firm. Column (4) distinguishes between innovations with technical aspects (product and process) and organizational changes, while column (5) distinguishes between product, process and organizational changes. Column (6) investigates industry heterogeneity, where "HT manuf." represents High-tech and Medium High tech manufacturing industries, "Manuf" other manufacturing, "HT services" high-tech knowledge intensive services and "Services" other service industries, all following Eurostat definitions. All specifications include skill-year (e.g. college-year) fixed effects, Mincer variables, firm fixed effects. Worker fixed effects are also included in Norway, while we apply the matching procedure for Hungary. 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: The Impact of the R&D Tax Credit Policy in Norway

	(1) College employment share	(2) College to non-college employment ratio	(3) Employment	(4) College premium	(5) College premium
Treatment effect	0.089***	0.104**	0.054	0.059**	0.031
	(0.031)	(0.047)	(0.060)	(0.028)	(0.031)
Worker FEs	N/A	N/A	N/A	No	Yes
Sample	Firm level	Firm level	Firm level	Worker level	Worker level
Observations	14,496	14,496	14,637	10,527,645	10,503,753
R-squared	0.94	0.96	0.91	0.21	0.41

Notes: This table shows how an R&D tax credit, introduced in 2002 in Norway, affected treated and control firms. Treated firms are those whose R&D expenditures had been below the policy threshold, NOK 4 mn, on average between 1999 and 2001. Control firms spent between NOK 4-12 mn in the same period. Columns (1), (2) and (3) report δ (the coefficients of the $Treat_j \times Post_t$) from the regression Equation (11), when the dependent variables are (log) college employment share (number of college workers divided by all workers), (log) college to non-college ratio, and (log) total employment, respectively. Columns (4) and (5) report δ^s (the coefficients of the $Treat_j \times Post_t \times College_i$) from the regression Equation (12), when the dependent variable is log wage. Column (4) includes skill-year (e.g. college-year) fixed effects, Mincer variables and firm fixed effects. Column (5) includes worker fixed effects as well. All regressions exclude the years 2002-2004 immediately following the reform and we restrict the sample to firms with at least 50 employees. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 8: The Contribution of Technological Change to Economy-wide College Premium over a 10-year Period

Form of Tech. Change	(1) Any	(2) R&D	(3) non R&D	(4) New	(5) Not new	(6) Only tech.	(7) Only org.	(8) Both
Panel A: Norway								
Reallocation effect	0.52%	0.52%	0.04%	0.34%	0.07%	0.01%	0.01%	0.37%
Wage premium effect	5.58%	4.93%	0.76%	3.49%	1.15%	0.40%	0.34%	4.11%
Total Effect ($\Delta\Theta$)	6.10%	5.44%	0.80%	3.83%	1.22%	0.42%	0.35%	4.48%
Panel B: Hungary								
Reallocation effect	3.74%	2.67%	2.04%	2.01%	3.57%	0.30%	0.19%	3.25%
Wage premium effect	10.09%	6.12%	6.44%	3.15%	10.44%	1.21%	0.47%	9.19%
Total Effect ($\Delta\Theta$)	13.83%	8.80%	8.49%	5.16%	14.00%	1.50%	0.66%	12.44%

Notes: This table shows the change in the aggregate college premium (in percentage points) due to firm-level technological change for a 10-year period based on Equation (15). The reallocation effect represents the change in wage premium resulting from workers moving between firms introducing new technology (innovative) and firms which do not do that (non-innovative firms). The wage premium effect captures the change in wage premium in firms introducing new technologies (innovative firms) and Total is the sum of the reallocation and wage premium effects, which reflects the overall contribution of technological change to inequality. The different columns quantify the contribution of firms conducting different forms of innovation to the aggregate college premium. We measure different forms of technological change from the detailed questionnaire of the CIS survey on firms' innovation activities. Column (1) captures the contribution of all innovative firms. Columns (2) and (3) calculate the contribution of innovators that conduct R&D and of those that do not, respectively. Columns (4) and (5) distinguish between innovators with new to the market innovations, and those whose innovations are only new to the firm. Finally, columns (6), (7) and (8) calculate the contributions of firms which conducted innovations only with technical aspects (product and process), only with organizational changes, or both, respectively.

Table 9: Economy-wide Skill Premium, Skill Ratio and Skill-Bias

	(1) $\Delta \ln \frac{H}{L}$	(2) $\Delta \ln \frac{w_H}{w_L}$	(3) $\Delta\Theta$	(4) Implied σ
Panel A: Norway (Change between 2005 and 2015)				
1) No skill bias	0.43	-0.09	0.00	4.88
2) With skill bias	0.43	-0.09	0.06	2.87
Panel B: Hungary (Change between 2000 and 2015)				
1) No skill bias	0.69	-0.07	0.00	9.35
2) With skill bias	0.69	-0.07	0.21	2.47
Panel C: United States (Change between 2000 and 2015)				
1) No skill bias	0.29	0.03	0.00	-9.8
2) With skill bias	0.29	0.03	0.09	4.75

Notes: This table shows the actual economy-wide change in (log) college to non-college ratio (column 1) and in (log) skill premium (column 2) for Norway (panel A), for Hungary (panel B), and for the United States (panel C). The country-level data come from the OECD Education at a Glance 2014 and 2020 data. For Norway the college premium is missing for 2000 and so we report the changes between 2005 and 2015. In Column (3) we explore various assumption on the extent to which technological change is skill biased (column 3). Then in column (4) we calculate the implied elasticity of substitution between college and non-college workers, σ, that is needed to explain the aggregate changes in the skill premium and skill ratio according to Equation (17). In each panel, row 1) assumes no skill-biased technological change, $\Delta\Theta$=0. In row 2), we apply our estimated total change in skill bias from column (1) of Table 8 after adjusting it to a 15-year period for Hungary and for the United States. For instance, for Hungary in Panel B of Table 8, the estimated total change in college premium due to technological change is 13.8% for a 10-year period, which implies 20.7% change for a 15-year period. For the United States we apply the estimated skill bias contribution from Norway, which is more similar to the Unite States in terms of its distance from the technology frontier.

Appendix A Additional Tables and Figures

A.1 Cross-country Relationship between Innovation and Skill Premium

Figure 1 in the main paper shows the cross-country relationship between the share of innovative firms and skill premium among "old" EU member states. Here we discuss the data and regression behind it.

We use country-level data from Eurostat's webpage on the premium of college educated workers, on the share of innovative firms, and on the share of firms conducting R&D activities. The source of R&D and innovation variables is the 2014 Community Innovation Survey (CIS) conducted in 23 (mainly EU) countries. Innovative firms are those which change their technology between 2012 and 2014 by introducing any new or significantly modified products/services/technologies/organizational solutions, which are new from the viewpoint of the firm. Therefore, innovation, according to this broad definition, does not have to be R&D-driven. The college premium data is calculated from the 2014 wave of the Structure of Earnings Survey.

We run cross-sectional regressions of the form:

$$\text{college premium}_j = \alpha + \delta_{inn}\text{ShareInnov}_j + \delta_{R\&D}\text{ShareR\&D}_j + \gamma X_j + \epsilon_j \tag{A.1}$$

where ShareInnov_j is the share of innovative firms in country j, $\text{ShareR\&D}_j$ is the share of R&D conducting firms, and X_j includes three variables: the share of college educated workers; CEE_j, which shows whether the country is a new member state (i.e. admitted after 2000); and log GDP per capita.

Table A.1 shows the estimates from this cross-sectional regressions. Column (1) shows that there is a positive and statistically significant (at the 5% level) relationship between the share of innovative firms and the college premium among old EU member states. Column (2) includes the new EU member states as well as controls for economic development (log GDP per capita) and the college share. The estimated relationship is almost the same, though the estimates become a bit noisy. Columns (3) and (4) show the estimates when we replace the share of innovative firms with the share of R&D-conducting ones. Surprisingly, no clear relationship emerges here, which underscores the key role non-R&D innovative firms play especially in countries farther from the technology frontier. Finally, in column (5) we include both the share of innovative and the share of R&D-conducting firms. We find that the share of innovative firms is more closely correlated with the college premium than the share of R&D-conducting firms, which again corroborates our key finding that non-R&D based innovation is responsible for a substantial amount of skill bias in technological change.

Table A.1: Innovation and the College Premium: Cross-country Evidence

LHS: College premium	(1)	(2)	(3)	(4)	(5)
Innovative firms (share)	0.894**	0.909*			0.832
	(0.408)	(0.486)			(0.606)
R&D firms (share)			-0.130	0.530	0.049
			(0.521)	(0.576)	(0.662)
Share of college educated		-0.013**		-0.017**	-0.015**
		(0.005)		(0.007)	(0.007)
GDP/capita		-0.001		0.206	0.043
		(0.186)		(0.180)	(0.211)
CEE		0.361**		0.303**	0.370**
		(0.129)		(0.130)	(0.136)
Constant	0.945***	1.443	1.490***	-0.203	1.058
	(0.237)	(1.764)	(0.135)	(1.788)	(1.970)
Sample	No CEE	All	No CEE	All	All
Observations	17	23	16	22	22
R-squared	0.242	0.479	0.004	0.433	0.493

Notes: This table shows the cross-country relationship between the college premium and the share of innovative firms (δ_{inn}) and the share of R&D-conducting firms ($\delta_{R\&D}$) from the regression equation A.1. Innovative firms are those firms changing their technology between 2012 and 2014 by introducing any new or significantly modified products/services/technologies/organization, which are new from the viewpoint of the firm, but that are not necessarily new to the market. Therefore, innovation, according to this broad definition, does not have to be R&D-driven. Columns (1) and (3) show the raw correlation among the old EU member states. Columns (2), (4) and (5) include all EU members states in the regression as well. CEE is a dummy for new EU member states. Standard errors in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

A.2 Country-industry Level Relationship between Technological Change and Skill Demand

In this section we complement our findings on firm-level technological change and skill demand and present evidence at the country-industry level. For this exercise, we use data from the Eurostat, which reports statistics on innovation activities as well as the share and premium of college educated workers at the 1-digit country-industry level. The source of innovation variables is the Community Innovation Survey (CIS). The college share and college premium is calculated from the Structure of Earnings Survey (SES). We have access to the micro data for both the CIS and SES in Norway and Hungary. For the other countries we only have access to aggregate statistics that can be downloaded from Eurostat's webpage.[47]

Figure A.1 shows the descriptive relationship between the share of innovative firms in 2010 and the change in the skill premium and the skill share between 2010 and 2014. We apply the same definition of innovation as in the main paper: innovative firms in 2010 are those that introduced any new or significantly modified product/service/process/organizational change, which is new from the viewpoint of the firm, but not necessary to the market, between 2008 and 2010. Therefore, innovative firms are those experiencing technological change. The figure shows that there is a clear positive relationship between the share of innovative firms (our measure of technological change) and the change in the skill premium and the change in the skill ratio.

We investigate the robustness of these relationships in Table A.2. We follow Machin & Van Reenen (1998) by regressing the 4-year change in skill demand on the share of innovative firms. In particular, we run regressions of the type:

$$\Delta y_{cst} = \delta_{inn} innov_{cst} + \delta_{R\&D} R\&D_{cst} + \gamma_y y_{cst} + \eta_c + \zeta_s + \epsilon_{cst} \tag{A.2}$$

where c indexes countries, s industries (1-digit) and t time periods. Δy_{cst} is the long difference, the change of y_{cst} between years t and $t+4$. η_c are country fixed effects, while ζ_s are industry fixed effects. $innov_{cst}$ is the share of innovative firms, while $R\&D_{cst}$ is the R&D intensity of the industry (the ratio of the total R&D expenditures and total the revenue of firms at the industry-country level). This long-difference regression removes differences in the level of the skill premium an the skill ratio at the country-industry level and identifies 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 weight the regressions by the number of firms in the CIS in the given country-industry cell to give more weight to observations which represent an average calculated from more observations. We cluster standard errors at the country level, because skill premia are likely to be strongly correlated within each country.

Table A.2 presents the regression results both for the change in the share of college educated workers (top panel) and the college premium (bottom panel). Column (1) reports basic regressions

[47] This merged sample includes EU27 countries (with the exception of Greece, Malta) and Norway, altogether 25 countries.

when both the share of innovative firms and the R&D intensity are included.[48] The estimates suggest that the increase in skill demand is linked to broadly defined innovation activities rather than only R&D. A 10 percentage point higher share of innovative firms is associated with a 1 percentage point stronger growth of the college employment share and a 3 percentage points higher increase in the college premium at the industry level. The coefficient of the R&D variable is small and often has a negative point estimate.

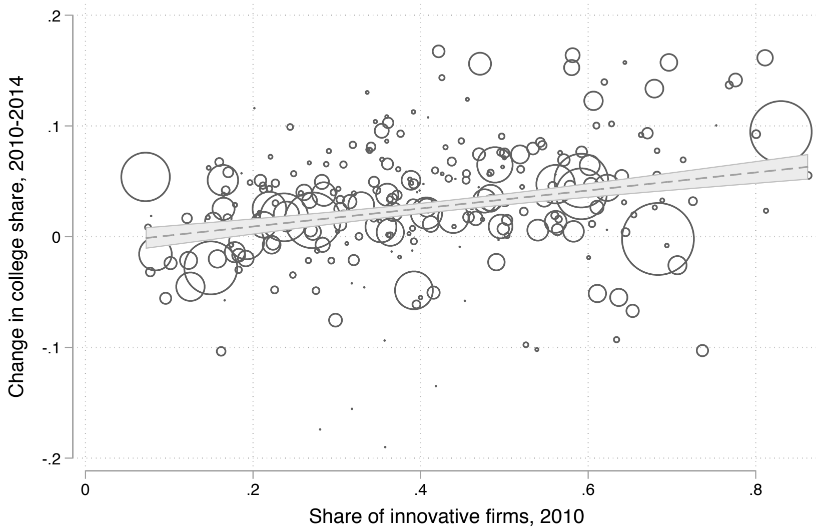
Column (2) includes country fixed effects to control for country-level shocks to skill supply or economic growth, while column (3) includes industry fixed effects, but not country fixed effects. The inclusion of these fixed effects has a small impact on the point estimates, even though some of the coefficients now are more nosily estimated. In Column (4) we include both country and industry fixed effects. The change in college share becomes insignificant here, while the point estimates of the college premium are unaffected by including both fixed effects. This highlights that there is a strong relationship between the share of innovative firms and subsequent increase in the college premium.

We can conclude from this exercise that our broadly defined innovation measure that captures many different forms of technological change (including technology adoption) is strongly related to skill demand at the country-industry level as well. For most specifications, we also see a response both in the relative quantity (college share) and in the relative wage margin (college premium), which motivates our investigation of both margins at the firm level.

[48] Including only the share of innovative firms in the regression yields similar results.

Figure A.1: Technological Change and the Change in Skill Demand: Country-industry Level Analysis

Panel A: Share of Innovators and the Change in the Share of College-educated Employees



Panel B: Share of Innovators and the Change in the College Premium

Notes: The figures illustrate the relationship between the share of innovative firms and subsequent change in skill demand at the 1-digit country-industry level for 25 European countries. We apply the same definition of innovation as in the main paper: innovative firms in 2010 are those that introduced any new or significantly modified product/service/process/organizational change, which is new from the viewpoint of the firm, but not necessary to the market, between 2007 and 2010. Therefore, innovative firms are those experiencing technological change. In particular, they show how the share of innovative firms in 2010 is related to the change in the share of college educated workers (Panel A) and the change in college premium (Panel B) between 2010 and 2014. The size of the circles is proportional to the number of firms in that cell, and the line shows a weighted regression line with a 95 percent confidence interval.

Table A.2: Technological Change and the Change in Skill Demand: Country-industry Level Regression Analysis

	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

Notes: These tables show the relationship between technological change and skill demand at the 1-digit country-industry level for 25 European countries. We present the estimated coefficients of the share of innovative firms (δ_{inn}) and R&D intensity ($\delta_{R\&D}$) from regression Equation (A.2). The dependent variable is the change in the share of college educated workers (top panel) and college premium (bottom panel). The main explanatory variables are the share of innovative firms and the industry's R&D intensity. We apply the same definition of innovation as in the main text: innovative firms in 2010 are those that introduced any new or significantly modified product/service/process/organizational change, which is new from the viewpoint of the firm, but not necessary to the market, between 2008 and 2010. All columns include the dependent variable in 2010. Column (2) also includes country fixed effects, column (3) industry fixed effects, while column (4) both country and industry fixed effects. Observations are weighted by the number of firms in the country-industry cell. Standard errors, clustered at the country level, are reported in parentheses. Significance levels are: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

A.3 Firm-level Technological Change and Change in the Structure of Earnings and Hours Worked

In the main paper we studied the relationship between change in wages and technological change. Here we investigate the relationship between various components of wages and hours worked and technological change by applying the same methodology – estimating Equation (9). Since the employer-employee register in Norway does not have detailed information on the structure of earnings, we will use the annual wage survey (the Norwegian version of the Structure of Earnings Survey) for this purpose. Table A.3 reports the change in various components. Column (1) shows the change in total salary following technological change. Column (2) shows the change in base wage without any bonus payments following innovation. The estimated change in skill premium is very similar with and without bonus payments both for Norway (2.1% with bonus payments vs. 1.9% without bonus payments) and for Hungary (6.7% with bonus payments vs. 7.8% without bonus payments). This highlights that the change in the skill premium is not driven by bonus payments rewarding the implementation of a successful innovation but rather reflect genuine technological change.

Column (3) of Table A.3 reports estimates when the outcome is working hours (instead of total salary) in regression Equation (9). We find no significant change in working hours of college workers (relative to non-college workers). Therefore, it is unlikely that the estimated effect on the wage premium results from longer hours worked by college workers after innovation.

Finally, in Norway, we can also assess whether whether non-cash benefits (taxable in-kind benefits reported in the employer-employee register) change following technological change. Column (4) in panel A reports the key estimates. We find no indication of changes in non-cash benefits for college workers (relative to non-college workers). Note that non-cash benefits can be interpreted as a proxy for the relative change in amenities. Nevertheless, we find no indication that this component of amenities changed in response to technological change.

Table A.3: Firm-level Technological Change and the Change in the Structure of Earnings in Hours Worked

Panel A: Norway

	(1) Total salary	(2) Base salary	(3) Hours	(4) Non-cash benefits
Innovation	-0.005	-0.004	0.001	-0.031
	(0.005)	(0.005)	(0.002)	(0.025)
Innovation x College	0.021**	0.019**	-0.000	0.010
	(0.009)	(0.008)	(0.001)	(0.021)
Observations in CIS	4,182,655	4,180,110	4,182,655	3,837,347
R-squared	0.73	0.72	0.58	0.83

Panel B: Hungary

	(1) Total salary	(2) Base salary	(3) Hours
Innovation	-0.008	-0.010	0.001
	(0.009)	(0.010)	(0.002)
Innovation x College	0.067***	0.078**	-0.002
	(0.023)	(0.033)	(0.002)
Observations in CIS	197,065	197,064	197,065
R-squared	0.70	0.71	0.70

Notes: This table shows robustness checks for the results on the change in workers' (log) wages following firm-level technological change presented in Table 2. We report the estimated coefficients on the innovation dummy, δ^u, and the innovation x college interaction, δ^s, from Equation (9), with different dependent variables. All specifications include skill-year (e.g. college-year) fixed effects, Mincer variables and firm fixed effects. Worker fixed effects are also included in Norway, while we apply the matching procedure for Hungary. Column (1) shows the change in total hourly wage, column (2) shows the change in base wage, while column (3) shows the change in working hours. The source of these variables in Norway (Panel A) is the Wage Survey rather than the administrative data used in the main regressions. Column (4) estimates the change in non-cash benefits. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

A.4 Firm-level Technological Change, Polarization, and Changes in Tasks

So far, we have classified workers into two skill groups, and looked at whether innovation affects the skill premium for college workers relative to non-college workers. However, Acemoglu & Autor (2011b) argue that the middle-skilled occupation categories, such as middle-skilled clerical, administrative, production and operative occupations, tend to be more affected by "routinization" than either high or low-skilled occupation categories, and that this has contributed to the observed wage polarization in the US. We study wage polarization across the skill distribution by interacting the innovation dummy in Equation (9) with the more detailed schooling variable which can take four values. The four groups are primary schooling, secondary schooling, vocational education, and college. Table A.4 reports results when omitting the primary schooling category. The coefficients of the interaction terms presented in the table hence show the changes in wages following an innovation relative to workers with the lowest education level. Note that the regressions still include interacted skill-year fixed effects. The results in the Table provide little evidence for wage polarization, neither in the cross-sectional specifications (columns (1) and (2)), nor in the specifications with firm fixed effects (columns (3) and (4)). The details differ slightly in the two countries: in Norway, workers with vocational training seem to benefit from innovation relative to workers with only primary or secondary education, while in Hungary the wages of the lower three educational categories do not seem to change after innovation takes place, while the wages of college educated workers increase substantially.

In the framework of this paper, we follow the seminal work by Katz & Murphy (1992) and Goldin & Katz (2010) and model technological change as potentially increasing the productivity of skilled workers (relative to the unskilled) in production. An alternative (or complementary) framework of technological change is a task-based one, where technological change affects both the productivity of high- and low-skilled labor in performing different tasks, as well as the allocation of tasks between the different types of labor (Autor et al. 2003, Acemoglu & Autor 2011b, Acemoglu & Restrepo 2020). Having a college degree may strongly be correlated with performing non-routine tasks, and our finding that innovation affects the skill share and the skill premium may capture changes in the task mix of firms, rather than solely the change in the productivity of performing different tasks.

To investigate this possibility, we create a measure of the degree to which an occupation contains routine tasks (RTI) following Autor et al. (2003).[49] Next, we include an interaction of $1 - RTI$ with the innovation dummy in regression Equation (9). A higher $1 - RTI$ represents a higher non-routine content of the worker's occupation. The results from this exercise are presented in Table A.5. We find that people working in less routine jobs are paid higher wages in general, even when controlling for worker fixed effects in Norway and estimating on the matched sample in Hungary. Firm-level technological change, however, does not affect this task content premium once we include person effects in the regression in Norway or apply the matching procedure in Hungary (see Column 4). Probably even more importantly, innovation's college premium is not affected by the inclusion of the task content variables, showing that the effect of innovation on the college premium does not only reflect the different task content of the jobs performed by college and non-college workers.

[49]We map the US occupation codes to Norwegian and Hungarian occupation codes.

Table A.4: Technological Change and the Change in Wages for Workers with Different Educational attainment

Panel A: Norway

	(1)	(2)	(3)	(4)
Innovation	0.104***	0.091***	-0.026**	-0.019**
	(0.016)	(0.014)	(0.011)	(0.009)
Innovation x Vocational	0.018	0.012	0.027*	0.041***
	(0.023)	(0.020)	(0.016)	(0.015)
Innovation x Secondary	-0.001	-0.002	0.001	0.009*
	(0.016)	(0.014)	(0.007)	(0.005)
Innovation x College	0.078**	0.068**	0.132***	0.053***
	(0.034)	(0.032)	(0.017)	(0.012)
Skill-year FE	Yes	Yes	Yes	Yes
Mincer variables	No	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes
Worker FEs	No	No	No	Yes
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373
R-squared	0.05	0.07	0.20	0.44

Panel B: Hungary

	(1)	(2)	(3)	(4)
Innovation	0.180***	0.133***	-0.040***	-0.003
	(0.019)	(0.016)	(0.013)	(0.015)
Innovation x Vocational	0.037**	0.049***	0.018**	0.005
	(0.017)	(0.014)	(0.008)	(0.012)
Innovation x Secondary	0.015	0.032	0.011	-0.019
	(0.035)	(0.030)	(0.011)	(0.019)
Innovation x College	0.107***	0.133***	0.135***	0.060**
	(0.031)	(0.029)	(0.016)	(0.026)
Skill-year FE	Yes	Yes	Yes	Yes
Mincer variables	No	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes
Matched sample	No	No	No	Yes
Observations in CIS	785,443	785,443	785,419	197,065
R-squared	0.44	0.51	0.71	0.70

Notes: This table investigates whether firm-level technological change is associated with the polarization of workers' wages by distinguishing between four education categories rather than only non-college/college. The interactions show innovative firms' premia for each education category relative to the premium of workers with a primary degree. All specifications include skill-year fixed effects, representing interactions of primary, secondary, vocational and college dummies with year dummies. Column (1) shows the estimates when including only skill-year (e.g. college-year) fixed effects in the regression. Columns (2)-(4) also include Mincer variables (gender, age, tenure, tenure squared, a dummy for new entrant in both countries and hours worked and a dummy for part-time employees in Hungary where part-time workers are also included in the sample). Columns (3)-(4) add firm fixed effects to the regression. Column (4) includes worker fixed effects in Norway and applies the matching procedure for Hungary (discussed in detail in Section 4.1 and Appendix Section B.5). Standard errors are clustered at the firm level and are reported in parentheses.

Table A.5: Technological Change and the Skill Premium: the Role of Routine Task Intensity

Panel A: Norway

	(1)	(2)	(3)	(4)
Innovation	0.091***	0.094***	-0.016	-0.002
	(0.019)	(0.019)	(0.015)	(0.012)
Non-routine	0.079***	0.055***	0.058***	-0.005
	(0.011)	(0.010)	(0.013)	(0.010)
Innovation x College	0.064***	0.057**	0.121***	0.041***
	(0.024)	(0.024)	(0.016)	(0.010)
Innovation x Non-routine	0.012	0.017*	0.006	0.007
	(0.010)	(0.010)	(0.015)	(0.009)
Skill-year FE	Yes	Yes	Yes	Yes
Mincer variables	No	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes
Worker FEs	No	No	No	Yes
Observations in CIS	4,804,373	4,804,373	4,804,373	4,804,373
R-squared	0.06	0.08	0.20	0.44

Panel B: Hungary

	(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	No	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes
Matched sample	No	No	No	Yes
Observations in CIS	784,732	784,732	784,732	157,638
R-squared	0.46	0.52	0.72	0.70

Notes: This table investigates whether firm-level technological change is associated with changes in workers' wage premium in non-routine jobs. We augment the estimates in columns (1)-(4) of Table 2 with the innovation dummy interacted with non-routine intensity. Non-routine intensity measures the degree to which an occupation contains non-routine tasks following Autor et al. (2003). All specifications include skill-year fixed effects, representing interactions of primary, secondary, vocational and college dummies with year dummies. Column 1 shows the estimates when including only skill-year (e.g. college-year) fixed effects in the regression. Columns (2)-(4) also include Mincer variables (gender, age, tenure, tenure squared, a dummy for new entrant in both countries and hours worked and a dummy for part-time employees in Hungary where part-time workers are also included in the sample). Columns (3)-(4) add firm fixed effects to the regression. Column (4) includes worker fixed effects in Norway and applies the matching procedure for Hungary (discussed in detail in Section 4.1 and Appendix Section B.5). Standard errors, clustered at the firm level are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

A.5 Technological Change and Changes in Skill Demand by Local-Area Firm Density

As we describe in the main text, our model predicts that whenever firms face a more elastic labor supply (β is larger), we expect a larger impact on the skill ratio, and a smaller impact on the skill premium for the same increase in skill-biasedness, $\Delta\theta$ (see Equations (7a) and (7b) in the main paper). In our model, the firm-level labor supply is tightly linked to the dispersion of idiosyncratic preferences for working at a particular firm. A key component of this dispersion is commuting distance, and so this dispersion is likely to be larger whenever workers live in lower density areas (or areas where the average distance between firms is larger). As a result, we compare responses in local areas with different levels of firm density. We summarized the key results in Figure 2, while here we provide more details underlying those results.

To this end, we extend the base worker- and firm-level regressions (Equations (9) and (10)) with an interaction of the innovation variable and the log density of the local area where the firm is located. We define density following Ciccone & Hall (1996) as the number of firms per square kilometer.

Table A.6 shows the firm-level results. The point estimate of the interaction is always positive in both countries, even though it is not always significant. This suggests that innovation leads to a larger increase in the skill ratio in denser areas. At the same time, in Table A.7 where worker-level results are shown, we find that the point estimate of the interaction term of the skill premium is negative. This suggest that in denser areas the changes in the skill premium are more muted.

In Figure 2 in the main paper we calculate the implied percent change in skill ratio and skill premium at the 10th and 90th percentiles of the density distribution.[50]

[50] The corresponding log densities are 0.85 and 3.8 in Norway while 0.06 and 5.79 in Hungary. The larger range in Hungary reflects that we apply smaller local areas there. We have 175 local areas in Hungary and 47 in Norway even though Norway's land area is four times bigger.

Table A.6: Change in the Skill Ratio Following Firm-level Technological Change by Local-Area Density

Panel A: Norway

	(1) College employment share	(2) College to non college employment ratio	(3) log employment
Density	0.004***	0.009***	0.004
	(0.001)	(0.003)	(0.006)
Innovation	0.002	-0.008	0.096***
	(0.005)	(0.013)	(0.026)
Innov x density	0.003*	0.013**	-0.016*
	(0.002)	(0.005)	(0.009)
Observations	18,204	17,785	24,931
R-squared	0.07	0.05	0.07

Panel B: Hungary

	(1) College employment share	(2) College to non college employment ratio	(3) log employment
Density	0.002	0.002	0.013*
	(0.002)	(0.004)	(0.007)
Innovation	0.012**	0.011	0.047**
	(0.006)	(0.010)	(0.020)
Innov x density	0.002	0.008*	-0.011
	(0.003)	(0.005)	(0.008)
Observations	2,152	2,124	2,147
R-squared	0.10	0.16	0.51

Notes: This table shows the relationship between firm-level technological change and subsequent 6-year change in firm-level college employment share (column 1), in college to non-college ratio (column 2), and log employment (column 3) by local-area firm density. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. We augment the regression Equation (10) by adding local-area density and an interaction term between local-area density and the innovation dummy. Local-area density is defined as the number of firms (over the sample period) divided by the size of the area (in square km). The "Innovation" dummy indicates whether the firm innovated according to the current CIS wave or any of the previous two waves. The other two explanatory variables in columns (1)-(2) are long differences of log capital stock and log value added. In each regression we include the lagged dependent variable preceding the baseline year and industry-year fixed effects. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A.7: Change in the Skill Premium Following Firm-level Technological Change by Local-Area Density

Panel A: Norway

	(1)
Innovation	-0.017
	(0.017)
Innovation x Log density	0.002
	(0.007)
College x Log density	0.097***
	(0.015)
Innovation x College	0.084***
	(0.021)
Innov x Log density x College	-0.016**
	(0.008)
Observations in CIS	4,804,373
R-squared	0.44

Panel B: Hungary

	(1)
Innovation	-0.015
	(0.014)
Innovation x Log density	0.003
	(0.005)
College x Log density	-0.005
	(0.009)
Innovation x College	0.080**
	(0.036)
Innov x Log density x College	-0.005
	(0.010)
Observations in CIS	195,627
R-squared	0.70

Notes: This table investigates the change in workers' (log) wages following firm-level technological change by local-area firm density. We augment the benchmark estimates reported in Column (4) of Table 2 by interacting local-area density with innovation, with college, and also add the triple interaction between local-area density, college and innovation. The regressions include skill-year fixed effects, Mincer variables and firm fixed effects. We also include worker fixed effects in Norway and apply the matching procedure for Hungary. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

A.6 Change in the Firm-level College Share Following Different Forms of Technological Change

In the main paper we discuss the relationship between various forms of technological change and the changes in the college premium. Here we present the changes in the college share. We investigate the heterogeneous effects of innovation by extending Equation (10). In particular, we include a set of indicator variables representing the different forms of technological change. The results are presented in Table A.8.

The first column reports the main estimate from column (1) of Table 5 as baseline. Column (2) investigates whether technological change by firms who conduct R&D is more skill-biased than non R&D-based changes. We study it by including both the basic innovation variable—capturing the effect of non-R&D innovation—and its interaction with an indicator variable showing whether the firm conducts R&D—capturing the additional effect of R&D-driven technological change. The regressions suggest that non-R&D innovation has a significant positive effect on college shares in both countries. Second, R&D-driven technological change seems to increase more the college shares than non-R&D innovation, though the differences are not statistically significant in Norway. Column (3) investigates whether it matters whether the innovation has a high novelty value, captured by an indicator variable measuring whether it is new for the market. The coefficients of this variable are small and insignificant, suggesting that 'new to the market' innovation is similar to other innovations.

Column (4) distinguishes between technical and organizational innovation. Note that a firm can conduct both at the same time; therefore, we introduce separate dummies for the different types of innovation. We find that both types lead to an increase in the college share. This reinforces the conclusions of Caroli & Van Reenen (2001) who found that organizational changes also increase the skill ratio. In column (6) we further distinguish between product and process innovation within innovations with technical aspects. There is only a minor difference between product and process innovation in both countries.

Finally, in column (6) we study whether the change in college share depends on the technology type of the sector. We classify industries into four groups: high- and low technology manufacturing, and high and low knowledge intensive services (see the details about the classification in footnote 38 in the main paper). In Norway, the change in college share is largest in high-tech manufacturing, and the lowest in high-knowledge (HK) services. In Hungary there seems to be a sharp contrast between manufacturing and services, with no evidence for changes in college share in services.

Figure 4 calculates the implied skill bias by combining the skill premium and skill ratio estimates in Equation 13.

Table A.8: Change in the College Share Following Different Forms of Technological Change

Panel A: Norway

	(1)	(2)	(3)	(4)	(5)	(6)
Innovation	0.010***	0.008***	0.008***			
	(0.002)	(0.003)	(0.003)			
Innovation x R&D		0.004	0.004			
		(0.003)	(0.003)			
Innovation x New			-0.000			
			(0.003)			
Technological				0.005**		
				(0.002)		
Organizational				0.011***	0.011***	
				(0.002)	(0.002)	
Process					0.002	
					(0.003)	
Product					0.005*	
					(0.003)	
Innovation x Manuf.						0.010***
						(0.003)
Innovation x HT manuf.						0.023***
						(0.009)
Innovation x Services						0.010***
						(0.003)
Innovation x HK services.						0.005
						(0.012)
Dependent variable (t-1)	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,205	18,205	18,205	18,205	18,205	18,205
R-squared	0.07	0.07	0.07	0.07	0.07	0.07

Panel B: Hungary

	(1)	(2)	(3)	(4)	(5)	(6)
Innovation	0.019***	0.011**	0.011**			
	(0.004)	(0.005)	(0.005)			
Innovation x R&D		0.013**	0.011*			
		(0.006)	(0.006)			
Innovation x New			0.019			
			(0.013)			
Technological				0.012***		
				(0.005)		
Organizational				0.010**	0.007	
				(0.005)	(0.007)	
Process					0.007	
					(0.007)	
Product					0.008	
					(0.005)	
Innovation x Manuf.						0.024***
						(0.006)
Innovation x HT Manuf.						0.018**
						(0.007)
Innovation x Services						-0.005
						(0.012)
Innovation x HK services						-0.022
						(0.034)
Dependent variable (t-1)	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,153	2,153	2,153	2,153	2,153	2,153
R-squared	0.10	0.10	0.10	0.10	0.10	0.10

Notes: This table shows the change in firm-level college employment shares following different forms of firm-level technological change. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. We extend the regression Equation 10. Column (1) reports the main estimate from column (1) of Table 5 as baseline. Column (2) includes a dummy showing whether the innovating firm conducted R&D and column (3) also includes a dummy showing whether the innovation was new for the firms' market rather than only for the firm. Column (4) distinguishes between innovations with technical aspects (product and process) and organizational changes, while column (5) distinguishes between product, process and organizational changes. Column (6) investigates industry heterogeneity, where "HT manuf." represents High-tech and Medium High tech manufacturing industries, "Manuf" other manufacturing, "HT services" high-tech knowledge intensive services and "Services" other service industries, all following Eurostat definitions. In each regression we include log capital stock, log value added, and the lagged dependent variable preceding the baseline year and industry-year fixed effects. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

A.7 Relationship Between the R&D Tax Credit Policy and Skill Demand

In Section 5.3 we study the effect of introducing an R&D tax credit scheme on skill demand. In this section we provide further details about the key results presented in Table 7 and also present some robustness checks.

The reform we study here is called Skattefunn and was introduced in 2002. The reform allowed firms to deduct 20 percent of their R&D expenditures up to a threshold of 4 million NOK. As a consequence of this, firms conducting R&D investments below the cost deduction threshold of 4 million NOK experienced a reduction in the marginal cost of investing in R&D. We therefore follow Bøler et al. (2015) and Bøler (2015) and classify a firm as treated if its pre-reform R&D investments are below 4 million NOK. More specifically, a firm will be considered treated if its average annual R&D investments in the years 1998-2001 are below 4 million NOK. We also restrict the sample to firms with at least 50 employees, for whom spending on R&D is likely.

The marginal cost of spending on R&D does not fall for firms spending more than the policy threshold on R&D, therefore the control group should be a subset of these firms. However, some of these firms are quite dissimilar from the treated group. Firms that spend substantially above the threshold are likely to be large, more globalised and more innovative than the treated group. If firm-level innovation tends to be skill-biased, heavily R&D-investing firms experience a continuous growth in the college employment share and premium relative to the treated firms even if the former group's R&D spending is not affected by the policy. In addition, thanks to the absolute nature of the threshold, non-treated firms tend to be larger, which may bias our firm-level estimates if small firms grow faster or change the structure of their labor force more rapidly. We therefore construct control groups in which firms spend above the policy threshold but below a certain percentile of the R&D expenditure distribution before the introduction of the policy. The choice of this threshold involves a trade-off: choosing a low value reduces the number of control firms, while a high threshold leads to the inclusion of firms which are very dissimilar from the treated firms into the control group. In our preferred control group firms spend between the policy threshold (4 million NOK, approx. 450,000 USD) and the median R&D spending in the distribution (12 million NOK), but we report sensitivity checks for this threshold.

We estimate the change in skill ratio following the introduction of the R&D tax credit policy using regression Equation (11) and the results are reported in columns (1) and (2) of Table 7. We estimate the effect of the introduction of the R&D tax credit on college premium using the regression Equation (12).

Table A.9 reports the results from Table 7 for different values of the threshold: 8 million NOK, 12 million NOK (the baseline) and 16 million NOK. The change in skill ratio (columns (1) and (2)) is similar for the different thresholds, with the point estimates increasing with the threshold value. The point estimates on the college premium are also similar across specifications, but the estimates become insignificant when the threshold is high.

Table A.9: The Impact of the R&D Tax Credit Policy in Norway: Applying Alternative Threshold Values for the Control Group

	(1) College employment share	(2) College to non-college employment ratio	(3) Employment	(4) College premium	(5) College premium
Panel A (Control firms 4-8 NOK mm)					
Treatment effect	0.068	0.070	0.032	0.063**	0.016
	(0.044)	(0.062)	(0.072)	(0.030)	(0.033)
Observations in R&D survey	13,025	13,025	13,025	2,398,437	2,398,437
R-squared	0.94	0.96	0.90	0.16	0.41
Panel A (Control firms 4-12 NOK mm)					
Treatment effect	0.076**	0.091*	0.048	0.059**	0.032
	(0.034)	(0.051)	(0.058)	(0.028)	(0.031)
Observations in R&D survey	13,359	13,359	13,359	2,568,739	2,568,739
R-squared	0.93	0.96	0.90	0.15	0.41
Panel A (Control firms 4-16 NOK mm)					
Treatment effect	0.086***	0.097**	0.091*	0.043	0.018
	(0.031)	(0.046)	(0.054)	(0.027)	(0.029)
Observations in R&D survey	13,569	13,569	13,569	2,638,158	2,638,158
R-squared	0.94	0.96	0.90	0.15	0.41
Sample	Firm level	Firm level	Firm level	Worker level	Worker level
Worker FEs	N/A	N/A	N/A	No	Yes

Notes: This table shows how an R&D tax credit, introduced in 2002 in Norway, affected treated and control firms. Treated firms are those whose R&D expenditures had been below the policy threshold, NOK 4 mn, on average between 1999 and 2001. Control firms spent between NOK 4-8 mn (in Panel A), between NOK 4-12 mn (Panel B) and between 4-16 mn (Panel C) in the same period. Columns (1), (2) and (3) report δ (the coefficients of the $Treat_j \times Post_t$) from the regression Equation (11) in the main paper, when the dependent variables are (log) college employment share (number of college workers divided by all workers), (log) college to non-college ratio, and (log) total employment, respectively. Columns (4) and (5) report δ^s (the coefficients of the $Treat_j \times Post_t \times College_i$) from the regression Equation (12), when the dependent variable is log wage. Column (4) includes skill-year (e.g. college-year) fixed effects, Mincer variables and firm fixed effects. Column (5) includes worker fixed effects as well. All regressions exclude the years 2002-2004 immediately following the reform and we restrict the sample to firms with at least 50 employees. Standard errors, clustered at the firm level, are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Appendix B Institutional Details and Data Appendix

B.1 Labor markets in Norway and Hungary

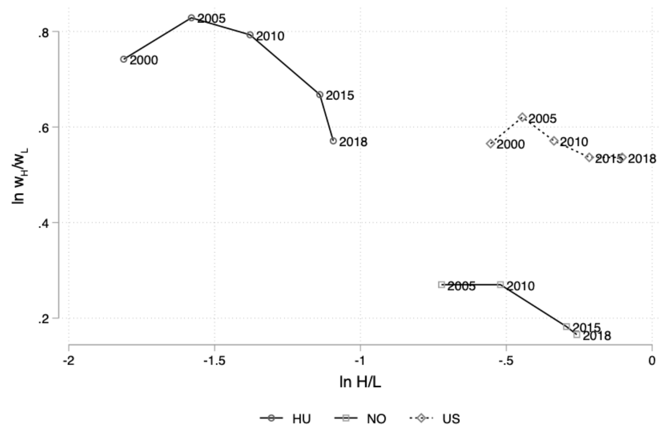
Norway's labor market is an example for the Nordic model, which has three key features: (i) flexible hiring and firing, (ii) a generous social safety net and (iii) active labor market policies. In the Nordic model, labor markets are less heavily regulated relative to other European labor markets, and collective agreements take over some of these functions. Unions density is very high in Norway, with more than 35% of workers in the private sector being Union members in 2012. Collective bargaining with the participation of Unions has led to smaller wage dispersion and sustained high wage growth (IMF 2015). Collective bargaining starts at the central and industry level, where key terms are decided, including a "floor" for wage increase. In the private sector, these central wage agreements are followed by firm-level collective bargaining. The firm-level wage agreements often lead to substantially higher wage increases and levels than the centrally agreed minimum wages, allowing for firm-level wage setting. For the majority of white-collar workers in the private sector, centrally negotiated collective agreements do not specify wages, therefore these workers have only firm-level wage formation, with strong individual-level elements (Nergaard 2014).

Hungarian employment protection institutions, in contrast, are closer to the Anglo-Saxon institutions than to those found in continental countries. It is relatively easy to dismiss workers (Tonin et al. 2009) and wage bargaining takes place mostly at the individual level. Collective wage bargaining is based on firm-level agreements with unions. Union membership was 10.2% percent in 2014, one of the lowest in the OECD.[51] Apart from firm-level bargaining, industry-level agreements are rare and set only weak requirements. Unions participate in the country-level bargaining forum called National Interest Reconciliation Council, which makes only non-binding recommendations (Rigó 2012). Employment contracts usually assume full time employment and pre-specify 8-hour working days. The actual working hours in these contracts are not monitored and firms can decide whether they want to measure and compensate for overtime hours. Part time work contracts add up to only 5 percent of the workforce and contracts on hourly basis are also rare.

Figure B.1 sketches the evolution of the two key variables in our study at the macro level for Norway, Hungary and the US between 2000 and 2018. The share of college graduates increased in all three countries throughout the period. This expansion started from a much lower level and was faster in relative terms in Hungary compared to the other two countries. In parallel with the education expansion, the skill premium fell in all three countries from 2005. The fall was strongest in Hungary, in line with the quick increase in the share of college workers. The evolution of the premium was nearly parallel in Norway and the US, but it is at a much lower level in Norway.

[51] OECD Employment and Labor Market Statistics.

Figure B.1: The evolution of the skill share and wage ratio in Norway, Hungary and the USA



Notes: H/L is based on the share of people with tertiary degrees among workers and the wage premium shows the average wage of 25-64 year-olds with income from employment compared to upper secondary education. Source: OECD Education at a Glance 2014, Table A62a and OECD Education at a Glance 2014 database ("eduadult" variable).

B.2 Innovation in Norway and Hungary

The European Innovation Scoreboard provides a comprehensive picture of innovation activities of European countries.[52] It uses four categories to rank the countries' innovation system, and classifies Norway as a 'Strong innovator' (the second group), and Hungary as a 'Moderate Innovator' (third group), suggesting that Norway is substantially closer to the world technology frontier than Hungary, where technology adoption plays a much larger role.[53]

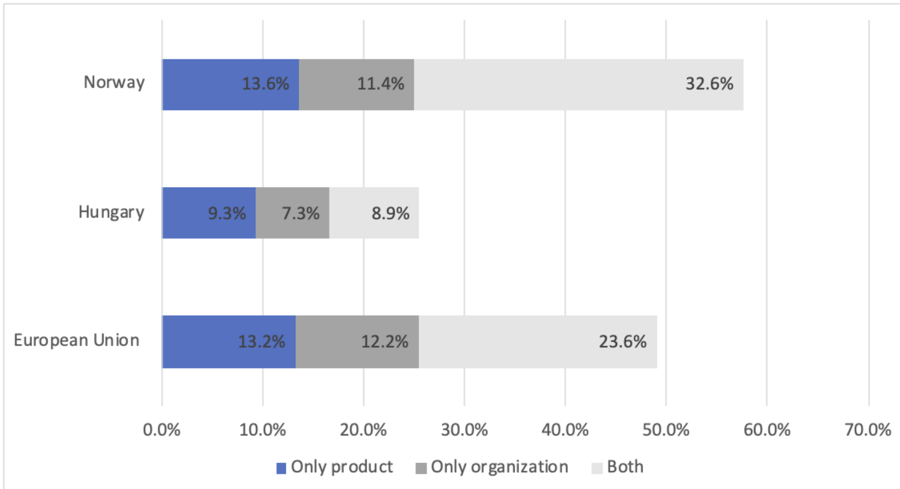
These differences are reflected by a number of indicators. In terms of GDP/capita, Norway's GDP was 20% above that of the USA (66 vs 55 thousand USD) and more than 150% above that of Hungary (25 thousand USD). On the innovation input side, the overall R&D/GDP ratio (in 2014) was 1.35% in Hungary and 1.71% in Norway compared to an EU average of 2% and 2.7% in the USA.[54]

Figure B.2 shows the share of firms conducting different types of innovation in the two countries and the average among the EU 27 and the United Kingdom. In Norway, 59% of firms are innovative compared to 25.5% in Hungary and 49% in the EU. Not only the share of innovative firms differs, but Norwegian innovators are much more likely to combine technical and organizational changes then either the EU or Hungary. Norwegian firms are also much more likely to rely on high novelty innovation while Hungarian firms conduct technology adoption to a larger extent. Among innovators, 26% introduced a "World first" innovation in Norway, compared to 5% in Hungary.

The CIS data also show characteristic differences in the inputs used by innovative firms in the two countries (Figure B.3). In line with a larger role of high-novelty innovations, Norwegian firms

[52]Available at https://ec.europa.eu/growth/industry/policy/innovation/scoreboards_en.
[53]We use numbers from 2014 around the end of our sample period, unless otherwise indicated.
[54]Source: https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm.

Figure B.2: Prevalence of Innovation Types in Norway and Hungary

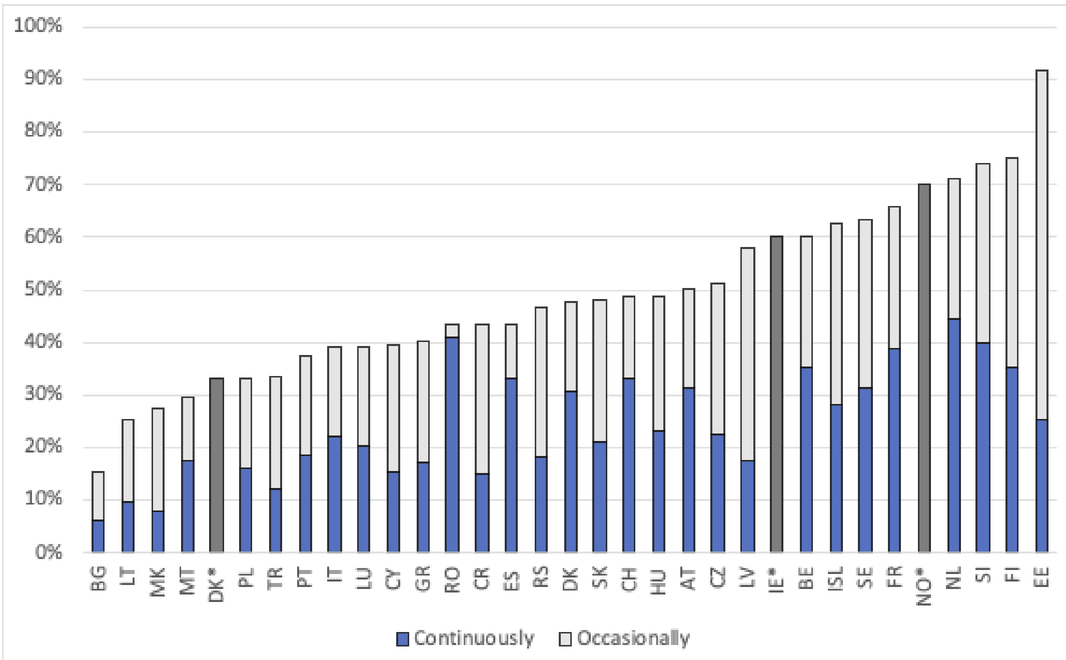


Notes: This Figure shows the share of innovative firms by the main type of innovation from the Community Innovation Survey 2014. "European Union" is the average of the EU 27 countries and the United Kingdom.

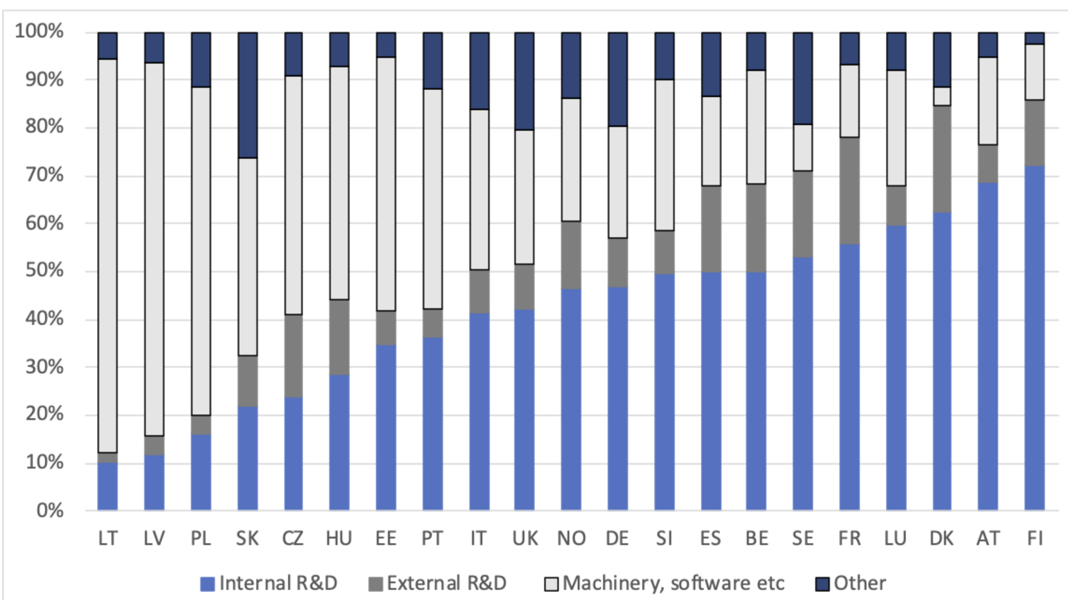
are much more likely to rely on R&D than Hungarian firms, with Norway having one of the highest share of R&D conducting firms among innovative firms (Panel A). Panel B shows a breakdown of the different types of innovation costs. It clearly demonstrates that the type of innovation costs captured by the CIS goes much beyond R&D, and also that in many European countries R&D is not the dominant component of innovation costs. The sum of external and internal R&D represents about 60% of Norwegian firms' innovation costs, but this number is closer to 45% in Hungary. In fact, the dominant innovation cost for Hungarian firms is machinery and software, in line with an innovation model which mainly relies on technology adoption, partly based on embodied knowledge (see e.g. Koren & Csillag 2017).

Figure B.3: Innovation inputs and outputs

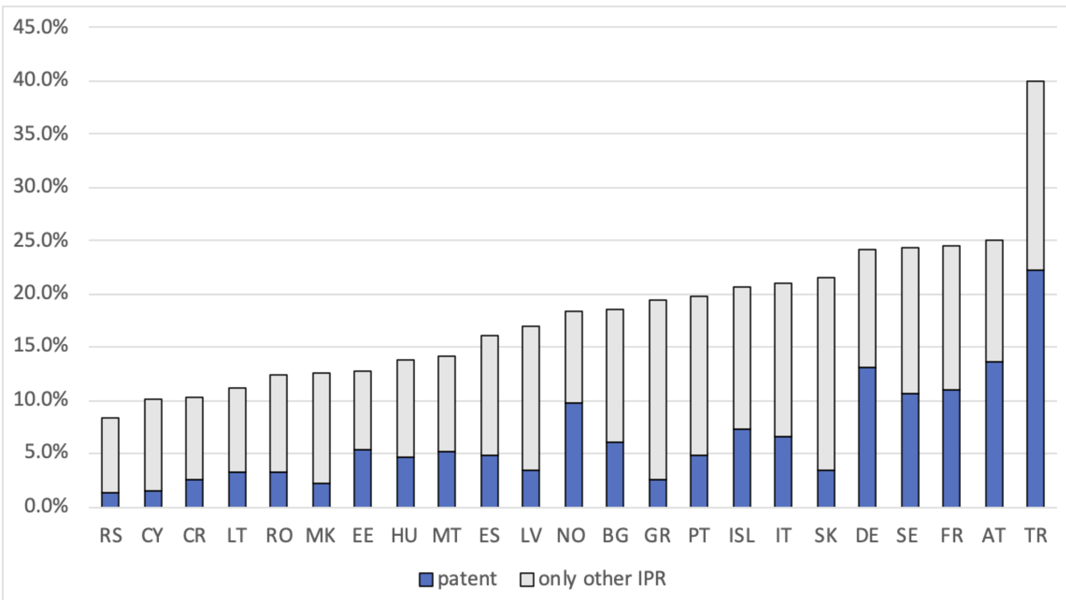
Panel A: Share of firms conducting R&D continuously or occasionally among technical innovators (%)



Panel B: Share of different expenditures in total innovation costs (%)



Panel C: Share of innovative firms applying for patents and other IP



Notes: Panel A of this figure shows the share of firms which conducted in-house R&D in firms which reported product and/or process innovations. The breakdown between continuous and occasional R&D spending is not available in countries denoted by *. Panel B shows the share of different types of innovation expenditures relative to total expenditures on product and process innovation. Panel C reports the share of firms which applied for a patent or other IP (a European utility model or that registered an industrial design right or a trademark). All the figures refer to the period between 2012 and 2014. Source: CIS, 2014.

B.3 Estimating Worker Fixed Effect in Norway

To estimate the impact of innovation on the skill premium, given by regression Equation (9), we use the sample of 4,804,373 worker-year observations for which we have information on innovation from the CIS (as described in Section 3). However, we make use of the full universe of workers in private sector firms (8,330,444 worker-year observations) to estimate the worker fixed effects (and other control variables) included in equation (9). More specifically, we "dummy out" the effect of innovation on the skill premium for observations for which innovation status is missing, such that only the 4,804,373 observations from the CIS contribute to identifying the effect of innovation on the skill premium. In the results tables, we report the observations numbers for which we have information from the CIS.

For the robustness results presented in Columns (3) and (4) of Table 3, we use data on occupations from the annual wage survey that covers only a subset of the workers in the main sample. We run these robustness regressions on the full sample of workers, but dummy out the effect of innovation on the skill premium for observations for which occupational status is missing.

B.4 Estimating Skill-Specific Firm Effects

We include skill-specific firm effects as follows. We group firms into deciles based on their college premium and then we include an additional interaction of firm premium-type deciles with the college dummy in the regression Equation (9). Ideally, we would like to group firms to skill premium deciles based on their residual college premium that is calculated after observed (X_{ijt}) and unobserved differences (η_i) are filtered out. This is not directly observed in the data and so we implement an iterative procedure. First, we group firms based on the average college premium over the sample period. Then we estimate Equation (9) and for each firm calculate the residual college premium conditional on X_{ijt} and on the estimated η_i. We then re-classify firms into college premium deciles based on the newly estimated residual college-premium. We re-estimate the model with new college-premium deciles and then we again calculate the residual college premium and re-classify firms into deciles based on that. This iterative procedure is repeated ten times when the college-premium deciles of most firms do not change any more at the reclassification. Including firm-skill fixed effects has a limited effect on the estimated change in wage premium. Also, whether we apply the 10 iteration or not we get very similar estimates.

B.5 The Matching Procedure in Hungary

The steps of the matching procedure in Hungary are the following. First, we run a probit regression with the innovation dummy as the dependent variable and basic firm characteristics as explanatory variables, while 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) when modeling the drivers of innovation at the firm level. 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 main market of the worker's firm is international, whether it received funding from the 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.

Based on this probit specification, we estimate a propensity score of innovating. 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 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. In our main specification we use a nearest neighbor matching, but results from kernel matching yield similar results. The matching procedure effectively excludes both frequent innovators and firms which are very unlikely to innovate, and we are hence more likely to compare quite similar firms. This presumption is reinforced by the fact that no pre-trend is detectable in this sample.

Appendix C Model

C.1 Basic Set-up

This section describes the firm's and worker's problem in detail. We also define the equilibrium and derive Equations (3c) and (3d) in the main paper. Throughout the section we drop the time subscript from the notation.

Worker's side. We model the worker's decision as in Card et al. (2018). For workers in skill group $S \in \{L, H\}$, the indirect utility of working at firm j is

$$u_{ij} = \ln \tau w_{Sj}^{\lambda} + \ln a_{Sj} + \epsilon_{ij} \tag{C.1}$$

where w_{Sj} is the firm-specific wage paid to individual i who belongs to skill group S, τ and λ approximate the progressivity of the tax system, and $\ln a_{Sj}$ is a firm-specific amenity common to all workers in group S, and ϵ_{ij} captures idiosyncratic preferences for working at firm j, arising e.g. from commuting distance, work flexibility and so on. We assume that the ϵ_{ij} are independent draws from a type I Extreme Value distribution with dispersion parameter ϕ.

Given posted wages workers are free to work at any firm they wish. Hence by standard arguments (McFadden et al. 1977) workers have logit choice densities of the following form:

$$\begin{aligned} P_{ij}^{s}\left(\arg\max_{k\in\{1,..J\}}\{u_{ij}^{s}\} = j\right) =& \frac{\exp\left(\frac{\lambda}{\phi}\ln w_{Sj} + \ln a_{Sj}\right)}{\sum_{k=1}^{J}\exp\left(\frac{\lambda}{\phi}\ln w_{Sk} + \ln a_{Sk}\right)} \\ =& \Lambda_S \exp\left(\frac{\lambda}{\phi}\ln w_{Sj} + \ln a_{Sj}\right) \end{aligned}$$

where $\Lambda_S = \frac{1}{\sum_{k=1}^{J}\exp\left(\frac{\lambda}{\phi}\ln w_{Sk}+\ln a_{Sk}\right)}$ is the same for all firms. This equation leads to the following upward sloping labor supply curve:

$$\ln S_j(w_{Sj}) = \ln\left(S \cdot P_{ij}^{s}\left(\arg\max_{k\in\{1,..J\}}\{u_{ij}^{s}\} = j\right)\right) = \ln(S\Lambda_S) + \beta \ln w_{Sj} + \ln a_{Sj}$$

where S is the total supply of workers from skill group S and $\beta = \frac{\lambda}{\phi}$.

Firm's side. Firms solve the following problem:

$$\pi_j(A_j, \theta_j) = \max_{w_{Hj}, w_{Lj}} p_j Q_j - H_j(w_{Hj})w_{Hj} - L_j(w_{Lj})w_{Lj} \tag{C.2}$$

Subject to

$$Q_j = A_j \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{C.3}$$

$$\ln p_j = \frac{1}{\rho} \ln \kappa_j - \frac{1}{\rho} \ln Q_j + \frac{\rho - 1}{\rho} \ln p + \frac{1}{\rho} \ln I \tag{C.4}$$

$$\ln L_j(w_{Lj}) = \ln(L\mathit{\Lambda}_L) + \beta \ln w_{Lj} + \ln a_{Lj} \tag{C.5}$$

$$\ln H_j(w_{Hj}) = \ln(H\mathit{\Lambda}_H) + \beta \ln w_{Hj} + \ln a_{Hj} \tag{C.6}$$

The first budget constraint (equation (C.3)) comes from the CES production function. While here we abstract away from capital or the presence of intermediate goods in the production function, we relax this assumption in Appendix Section C.3. The presence of capital does not change any of our conclusions presented here. The second budget constraint (Equation (C.4)) represents the firm-specific output demand function that firms face. We micro found this equation in Appendix Section C.4 using a monopolistic competition model and show that κ_j is a firm specific demand shifter, p is the price index and I is the total income of the consumer. The third (Equation (C.5)) and fourth (Equation (C.6)) budget constraints represent the upward sloping labor supply function we just derived above. As we describe above, $\mathit{\Lambda}_L$ and $\mathit{\Lambda}_H$ are determined by other firms' wage-setting behavior. Following Card et al. (2018) and Lamadon et al. (2018) we assume that firm's behavior has no direct effect on this outcome.

Equilibrium. We define the market equilibrium in the following way.

Definition 1. Given firm characteristics (A_j, θ_j, κ_j, a_{Hj}, a_{Lj}), worker distribution (L, H), and preference parameter (ϕ), we define equilibrium as the worker's decision on which firm to choose $j(i,t)$, market-level price index p, wage indices $\mathit{\Lambda}_H$, $\mathit{\Lambda}_L$, and firm's decision on prices p_j and wages w_{Hj}, w_{Lj} such that:

1. Workers choose firms that maximize their utility, as defined in equation (C.1).

2. Firms choose labor demand by setting wages w_{Sj} for each worker type to maximize profits (Equation (C.2)) subject to the production function (Equation (C.3)), product market constraint (Equation (C.4)) and labor supply constraints (Equations (C.5) and (C.6)). The market-level wage indices $\mathit{\Lambda}_L$ and $\mathit{\Lambda}_H$ are generated from the workers' optimal decisions $j(i,t)$, as described in Equation (C.1).

3. The market level wage indices ($\mathit{\Lambda}_L$ and $\mathit{\Lambda}_H$) and price index (p) are generated from the firms' optimal decisions on w_{Sj} and p_{Sj}.

Solution. We solve the firm problem described above.

The FOC of the problem is the following:

$$\frac{\partial \pi_j(A_j,\theta_j)}{\partial w_{Lj}} = Q_j \frac{\partial p_j}{\partial Q_j}\frac{\partial Q_j}{\partial L_j}\frac{\partial L_j}{\partial w_{Lj}} + p_j \frac{\partial Q_j}{\partial L_j}\frac{\partial L_j}{\partial w_{Lj}} - \frac{\partial L_j}{\partial w_{Lj}} w_{Lj} - L_j = 0 \tag{C.7}$$

$$\frac{\partial \pi_j(A_j,\theta_j)}{\partial w_{Hj}} = Q_j \frac{\partial p_j}{\partial Q_j}\frac{\partial Q_j}{\partial H_j}\frac{\partial H_j}{\partial w_{Hj}} + p_j \frac{\partial Q_j}{\partial H_j}\frac{\partial H_j}{\partial w_{Hj}} - \frac{\partial H_j}{\partial w_{Hj}} w_{Hj} - H_j = 0 \tag{C.8}$$

The first FOC, representing the decision about low-skilled workers, can be rewritten as

$$\left(\frac{Q_j}{p_j}\frac{\partial p_j}{\partial Q_j} + 1\right) p_j \frac{\partial Q_j}{\partial L_j}\frac{L_j}{w_{Lj}}\frac{\partial L_j}{\partial w_{Lj}}\frac{w_{Lj}}{L_j} - \left(1 + \frac{\partial L_j}{\partial w_{Lj}}\frac{w_{Lj}}{L_j}\right) L_j = 0$$

The second (Equation (C.4)) and third (Equation (C.5)) budget constraints imply that

$$\frac{1+\rho}{\rho} p_j \frac{\partial Q_j}{\partial L_j} = \frac{1+\beta}{\beta} w_{Lj}$$

The CES production function implies that:

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

and so we get the following expression for the FOC:

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\sigma-1}{\sigma}} Q_j^{\frac{1}{\sigma}} (1-\theta_j) L_j^{-\frac{1}{\sigma}} \frac{\beta}{1+\beta} = w_{Lj}$$

A similar expression leads to the following expression for high skilled workers:

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\sigma-1}{\sigma}} Q_j^{\frac{1}{\sigma}} \theta_j H_j^{-\frac{1}{\sigma}} \frac{\beta}{1+\beta} = w_{Hj}$$

Dividing the two first order conditions leads to the following expression

$$\frac{\theta_j H_j^{-\frac{1}{\sigma}}}{(1-\theta_j) L_j^{-\frac{1}{\sigma}}} = \frac{w_{Hj}}{w_{Lj}}$$

This can be rearranged to derive the relationship between the skill premium and the skill ratio, which will be the following:

$$\ln \frac{w_{Hj}}{w_{Lj}} = \ln \frac{\theta_j}{1-\theta_j} - \frac{1}{\sigma} \ln \frac{H_j}{L_j} \tag{C.9}$$

The second and the third budget constraints also imply that

$$\ln \frac{H_j}{L_j} = \ln \frac{H\Lambda_H}{L\Lambda_H} + \beta \ln \frac{w_{Hj}}{w_{Lj}} + \ln \frac{a_{Hj}}{a_{Lj}}$$

which leads to Equation (6a) in the main paper

$$\ln \frac{w_{Hj}}{w_{Lj}} = \frac{\sigma}{\sigma+\beta} \ln \frac{\theta_j}{1-\theta_j} - \frac{1}{\sigma+\beta} \ln \frac{H\Lambda_H}{L\Lambda_L} - \frac{1}{\sigma+\beta} \ln \frac{a_{Hj}}{a_{Lj}}$$

and to Equation (6b) in the main paper

$$\ln \frac{H_j}{L_j} = \frac{\beta\sigma}{\sigma+\beta} \ln \frac{\theta_j}{1-\theta_j} + \frac{\sigma}{\sigma+\beta} \ln \frac{H\Lambda_H}{L\Lambda_L} + \frac{\sigma}{\sigma+\beta} \ln \frac{a_{Hj}}{a_{Lj}}.$$

The relationship between the skill premium and the skill ratio can be also used to demonstrate the key idea of the paper. The change in skill premium in response to innovation will be the following:

$$\underbrace{\triangle \ln \frac{w_{Hj}}{w_{Lj}}}_{\substack{\text{Change in}\\ \text{skill premium}}} = \underbrace{\triangle \ln \frac{\theta_j}{1-\theta_j}}_{\substack{\text{Change in}\\ \text{skill bias}}} - \frac{1}{\sigma} \underbrace{\triangle \ln \frac{H_j}{L_j}}_{\substack{\text{Change in}\\ \text{skill ratio}}} \tag{C.10}$$

Since $\sigma \geq 0$, this equation shows that the skill premium (w_{Hj}/w_{Lj}) and skill ratio (H_j/L_j) will be negatively related when there is no change in the skill bias component. As a result, a joint increase in the premium and the skill ratio provides *prima facie* evidence for innovation activities being skill-biased.

C.2 Skill-Specific Dispersion in Idiosyncratic Preferences

Now we can extend the baseline framework by allowing differential dispersion of the idiosyncratic error term (ϵ_{ij}) for high- (ϕ_H) and low-skilled workers (ϕ_L). The upward-sloping labor supply curves firms face will have differential elasticities:

$$\ln S_j(w_{Sj}) = \ln(S\Lambda_S) + \beta_S \ln w_{Sj} + \ln a_{Sj}$$

where $\beta_S = \frac{\lambda}{\phi_S}$.

Solution. We follow the same steps as above. The FOC of the problem leads to the following two equations:

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\sigma-1}{\sigma}} Q_j^{\frac{1}{\sigma}} (1-\theta_j) L_j^{-\frac{1}{\sigma}} \frac{\beta_L}{1+\beta_L} = w_{Lj}$$

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\sigma-1}{\sigma}} Q_j^{\frac{1}{\sigma}} \theta_j H_j^{-\frac{1}{\sigma}} \frac{\beta_H}{1+\beta_H} = w_{Hj}$$

The ratio of the two first order conditions leads to the following expression

$$\frac{\theta_j H_j^{-\frac{1}{\sigma}} \frac{\beta_H}{1+\beta_H}}{(1-\theta_j) L_j^{-\frac{1}{\sigma}} \frac{\beta_L}{1+\beta_L}} = \frac{w_{Hj}}{w_{Lj}}$$

This can be rearranged to get the following relationship between skill premium and skill demand:

$$\underbrace{\ln \frac{w_{Hj}}{w_{Lj}}}_{\substack{\text{skill}\\ \text{premium}}} = \underbrace{\ln \frac{1+\frac{1}{\beta_L}}{1+\frac{1}{\beta_H}}}_{\substack{\text{relative}\\ \text{mark-down}}} + \underbrace{\ln \frac{\theta_j}{1-\theta_j}}_{\substack{\text{skill}\\ \text{bias}}} - \frac{1}{\sigma} \underbrace{\ln \frac{H_j}{L_j}}_{\substack{\text{skill}\\ \text{ratio}}} \tag{C.11}$$

The main difference between this equation and the one derived under constant dispersion (Equation (C.9)) is the new term reflecting the relative mark-down on the two labor markets. This new term reflects that the wage premium in this case also depends on the extent to which firm-level labor supply elasticities differ across skill groups. Nevertheless, it is worth pointing out that once we look at the change in skill premium and skill ratio, this mark-down term will cancel out as β_H and β_L are determined entirely by workers' preferences (i.e. the low and high skilled workers' dispersion of idiosyncratic preferences toward the workplace), which are unlikely to be affected by firm-level innovation activities. Thus, the change in skill premium will be driven by the following equation:

$$\underbrace{\triangle \ln \frac{w_{Hj}}{w_{Lj}}}_{\substack{\text{Change in}\\ \text{skill premium}}} = \underbrace{\triangle \ln \frac{\theta_j}{1-\theta_j}}_{\substack{\text{Change in}\\ \text{skill bias}}} - \frac{1}{\sigma} \underbrace{\triangle \ln \frac{H_j}{L_j}}_{\substack{\text{Change in}\\ \text{skill ratio}}} \tag{C.12}$$

This equation is the same as Equation (C.10), which was derived when $\beta_H = \beta_L$.

Going back to the problem of finding the equilibrium H_j and L_j, the above Equation (C.11) expresses the relationship between skill premium and skill demand. Then $\beta_L \neq \beta_H$, the third (Equation (C.5)) and fourth (Equation (C.6)) budget constraints become:

$$\ln L_j(w_{Lj}) = \ln(L\Lambda_L) + \beta_L \ln w_{Lj} + \ln a_{Lj} \tag{C.13}$$

$$\ln H_j(w_{Hj}) = \ln(H\Lambda_H) + \beta_H \ln w_{Hj} + \ln a_{Hj} \tag{C.14}$$

which implies that

$$\ln\frac{H_j}{L_j}=\ln\frac{H\Lambda_H}{L\Lambda_L}+\beta_H\ln w_{Hj}-\beta_L\ln w_{Lj}+\ln\frac{a_{Hj}}{a_{Lj}} \tag{C.15}$$

Unfortunately, we cannot express the solution simply in terms of the ratios of $\ln\frac{w_{Hj}}{w_{Lj}}$ and $\ln\frac{H_j}{L_j}$ as the solution will also depend on $\ln w_{Lj}$. While this latter can be expressed from one of the first order conditions, it is not possible to express the ratios in closed-form any more. Nevertheless, we can characterize the impact of changes of various factors on $\ln\frac{w_{Hj}}{w_{Lj}}$ and $\ln\frac{H_j}{L_j}$. We do this in Proposition 1.

Proposition 1. *Suppose firms maximize profits given the budget constraints in Equations (C.3), (C.4), (C.13), (C.14). Changes in A_j and κ_j have the following effect on the firm-level skill ratio $\left(\ln\frac{H_j}{L_j}\right)$ and on the wage ratio $\left(\ln\frac{w_{Hj}}{w_{Lj}}\right)$.*

1. *If $\beta_H=\beta_L$, then $\ln\frac{w_{Hj}}{w_{Lj}}$ and $\ln\frac{H_j}{L_j}$ are unaffacted by A_j and κ_j.*
2. *If $\beta_H>\beta_L$, then $\ln\frac{w_{Hj}}{w_{Lj}}$ is decreasing and $\ln\frac{H_j}{L_j}$ is increasing in A_j and in κ_j.*
3. *If $\beta_H<\beta_L$, then $\ln\frac{w_{Hj}}{w_{Lj}}$ is increasing and $\ln\frac{H_j}{L_j}$ is decreasing in A_j and in κ_j.*

Proof. We prove the proposition for A_j, but applying the same steps one can prove the statement for κ_j. Plugging Equation (C.15) into Equation (C.11) on the skill ratio leads to the following expression:

$$\sigma\left(\ln\frac{1+\frac{1}{\beta_L}}{1+\frac{1}{\beta_H}}+\ln\frac{\theta_j}{1-\theta_j}-\ln\frac{w_{Hj}}{w_{Lj}}\right)=\ln\frac{H\Lambda_H}{L\Lambda_L}+\beta_H\ln w_{Hj}-\beta_L\ln w_{Lj}+\ln\frac{a_{Hj}}{a_{Lj}} \tag{C.16}$$

Taking the derivative of that with respect to $\ln A_j$ leads to the following expression:

$$\sigma\frac{\partial\ln w_{Hj}}{\partial\ln A_j}-\sigma\frac{\partial\ln w_{Lj}}{\partial\ln A_j}=\beta_L\frac{\partial\ln w_{Lj}}{\partial\ln A_j}-\beta_H\frac{\partial\ln w_{Hj}}{\partial\ln A_j}$$

$$(\sigma+\beta_H)\frac{\partial\ln w_{Hj}}{\partial\ln A_j}=(\sigma+\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln A_j} \tag{C.17}$$

Since the third (Equation (C.13)) and the fourth (Equation (C.14)) budget constraints imply that $\frac{\partial\ln w_{Hj}}{\partial\ln H_j}=\frac{1}{\beta_H}$ and $\frac{\partial\ln w_{Lj}}{\partial\ln L_j}=\frac{1}{\beta_L}$, we can express $\frac{\partial\ln w_{Hj}}{\partial\ln A_j}$ and $\frac{\partial\ln w_{Lj}}{\partial\ln A_j}$ as

$$\frac{\partial\ln w_{Hj}}{\partial\ln A_j}=\frac{\partial\ln w_{Hj}}{\partial\ln H_j}\frac{\partial\ln H_j}{\partial\ln A_j}=\frac{1}{\beta_H}\frac{\partial\ln H_j}{\partial\ln A_j}$$

and

$$\frac{\partial\ln w_{Lj}}{\partial\ln A_j}=\frac{\partial\ln w_{Lj}}{\partial\ln L_j}\frac{\partial\ln L_j}{\partial\ln A_j}=\frac{1}{\beta_L}\frac{\partial\ln L_j}{\partial\ln A_j}.$$

Plugging these two expressions into Equation (C.17) leads to

$$\left(\frac{\sigma}{\beta_H}+1\right)\frac{\partial \ln H_j}{\partial \ln A_j}=\left(\frac{\sigma}{\beta_L}+1\right)\frac{\partial \ln L_j}{\partial \ln A_j} \tag{C.18}$$

It is easy to see that if $\beta_H > \beta_L$, then we have $\frac{\partial \ln H_j}{\partial \ln A_j} > \frac{\partial \ln L_j}{\partial \ln A_j}$ and $\frac{\partial \ln w_{Hj}}{\partial \ln A_j} < \frac{\partial \ln w_{Lj}}{\partial \ln A_j}$, and so

$$\frac{\partial \ln \frac{H_j}{L_j}}{\partial \ln A_j} > 0 \quad \textit{and} \quad \frac{\partial \ln \frac{w_{Hj}}{w_{Lj}}}{\partial \ln A_j} < 0.$$

$\square$

Proposition 1 states that the Hicks-neutral technological shock (A_j) or firm specific demand shifter (κ_j) directly affect the skill ratio and the skill premium if $\beta_H \neq \beta_L$. Nevertheless, the effects of these shocks on $\ln \frac{w_{Hj}}{w_{Lj}}$ and $\ln \frac{H_j}{L_j}$ always have a different sign. So if one of them increases, then the other will fall. This implies that demand shifters (κ_j) or Hicks-neutral shocks (A_j) cannot explain a joint increase in skill demand and skill ratio even if $\beta_H \neq \beta_L$.

Why does even a Hicks-neutral change (A_{jt}) affect the skill ratio when $\beta_H \neq \beta_L$? When a firm experiences an increase in A_{jt}, it will expand and, therefore, increase its demand for both type of workers. If, for example, $\beta_H > \beta_L$, high skilled workers are more responsive to changes in wages than the low skilled ones, and, therefore, firms can expand their skilled labor force more when they increase the wages of both types similarly. In optimum, firms adjust both on the wage and quantity margins: they raise high skilled workers' wages less ($\triangle \ln \frac{w_{hj}}{w_{lj}} < 0$), but hire more of them ($\triangle \ln \frac{h_j}{l_j} > 0$).

An important implication of Proposition 1 is that finding that the skill ratio is increasing after an innovation does not prove that the innovation is skill-biased. In the presence of non-competitive labor markets even an (exogenous) Hicks-neutral shock can affect the skill ratio if firms have different wage-setting power at the high and low skilled labor markets (for instance, if $\beta_H > \beta_L$). Nevertheless, as Equation (C.12) above demonstrated, whenever both the skill premium and skill ratio increases, we can conclude that technological change is skill-biased.

Now we also characterize how changes in the key parameters of the firm-level labor supply affect firm's behavior.

Proposition 2. *Suppose firms maximize profits given the budget constraints in Equations (C.3), (C.4), (C.13), (C.14). Then the change in $X = \{H\Lambda_H, a_{Hj}\}$ has the following impact on the skill premium and skill ratio*

$$\frac{\partial \ln \frac{w_{Hj}}{w_{Lj}}}{\partial \ln X} = -\left(\frac{1}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln X}\right)$$

$$\frac{\partial \ln \frac{H_j}{L_j}}{\partial \ln X} = \left(\frac{\sigma}{\sigma + \beta_H}\right)\left(1 + (\beta_H - \beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln X}\right)$$

where

$$1 + (\beta_H - \beta_L)\frac{\partial \ln w_{lj}}{\partial \ln X} = \frac{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}} A_j^{-\frac{1}{\sigma}} Q_j^{\frac{1}{\sigma}} - \beta_L}{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}} A_j^{-\frac{1}{\sigma}} Q_j^{\frac{1}{\sigma}} - \beta_L - \left(\frac{\theta_j H_j^{\frac{\sigma-1}{\sigma}}}{\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}}}\left(\frac{\sigma(\beta_H - \beta_L)}{\sigma + \beta_H}\right)\right)}$$

Proof. We prove the statement for $H\Lambda_H$, but the same steps could be used to prove the statement for a_{Hj}. As we derived in the proof of Proposition 1, the third (equation C.13) and fourth (Equation (C.14)) budget constraints together with the FOC (Equation (C.11)) imply that

$$\sigma\left(\ln \frac{\beta_H}{1+\beta_H} - \ln \frac{\beta_L}{1+\beta_L} + \ln \frac{\theta_j}{1-\theta_j} - \ln \frac{w_{Hj}}{w_{Lj}}\right) = \ln \frac{H\Lambda_H}{L\Lambda_L} + \beta_H \ln w_{Hj} - \beta_L \ln w_{Lj} + \ln \frac{a_{Hj}}{a_{Lj}}$$

Taking the derivative of that with respect to ln $H\Lambda_H$ leads to the following expression:

$$-\sigma\left(\frac{\partial \ln w_{Hj}}{\partial \ln H\Lambda_H} - \frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right) = 1 + \beta_H \frac{\partial \ln w_{Hj}}{\partial \ln H\Lambda_H} - \beta_L \frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}$$

This can be rearranged to

$$\frac{\partial \ln w_{Hj} - \partial \ln w_{Lj}}{\partial \ln H\Lambda_H} = -\frac{1}{\sigma + \beta_H}\left(1 + (\beta_H - \beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right)$$

Using that $\frac{\partial \ln w_{Hj}}{\partial \ln H\Lambda_H} = \frac{1}{\beta_H}\frac{\partial \ln H_j}{\partial \ln H\Lambda_H} - \frac{1}{\beta_H}$ and $\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H} = \frac{1}{\beta_L}\frac{\partial \ln L_j}{\partial \ln H\Lambda_H}$ from the budget constraints, one can also express the relationship between changes in wages as

$$\frac{1}{\beta_H}\frac{\partial \ln H_j}{\partial \ln H\Lambda_H} - \frac{1}{\beta_H} - \frac{1}{\beta_L}\frac{\partial \ln L_j}{\partial \ln H\Lambda_H} = -\frac{1}{\sigma + \beta_H}\left(1 + (\beta_H - \beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right)$$

$$\frac{\partial \ln H_j}{\partial \ln H\Lambda_H} - \frac{\partial \ln L_j}{\partial \ln H\Lambda_H} = \left(\frac{\sigma}{\sigma + \beta_H}\right)\left(1 + (\beta_H - \beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right)$$

Which proves the statement. Now we need to obtain the expression for $1 + (\beta_H - \beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}$. The FOC for low-skilled workers of the profit maximization problem implies that

$$\frac{1+\rho}{\rho}\left(\frac{I\kappa_j}{p^{1-\rho}}\right)^{\frac{1}{\rho}} A_j^{\frac{\sigma-1}{\sigma}} Q_j^{\frac{1}{\sigma}-\frac{1}{\rho}} (1-\theta_j) L_j^{-\frac{1}{\sigma}} \frac{\beta_L}{1+\beta_L} = w_{Lj} = \left(\frac{L_j}{L\Lambda_L a_{Lj}}\right)^{\frac{1}{\beta_L}}$$

We take the log:

$$\ln\frac{1+\rho}{\rho}+\ln\left(\frac{I\kappa_j}{p^{1-\rho}}\right)^{\frac{1}{\rho}}+\frac{\sigma-1}{\sigma}\ln A_j+\left(\frac{1}{\sigma}-\frac{1}{\rho}\right)\ln Q_j+\ln(1-\theta_j)-\frac{1}{\sigma}\ln L_j+\ln\frac{\beta_L}{1+\beta_L}=\ln w_{Lj}$$

and take the derivative with respect to $\ln H\mathit{\Lambda}_H$, which leads to

$$\left(\frac{1}{\sigma}-\frac{1}{\rho}\right)\frac{\partial\ln Q_j}{\partial\ln H\mathit{\Lambda}_H}-\frac{1}{\sigma}\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}=\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}.$$

Using that $\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}=\beta_L\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}$ (see Equation (C.15)) we get

$$\left(\frac{1}{\sigma}-\frac{1}{\rho}\right)\frac{\partial\ln Q_j}{\partial\ln H\mathit{\Lambda}_H}=\left(1+\frac{\beta_L}{\sigma}\right)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}$$

or

$$\frac{\partial\ln Q_j}{\partial\ln H\mathit{\Lambda}_H}=\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}.$$

Denoting $N_j=\left[\theta_jH_j^{\frac{\sigma-1}{\sigma}}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right]$ for notational convenience, we notice that

$$\begin{aligned}\frac{\partial\ln Q_j}{\partial\ln H\mathit{\Lambda}_H}=&\frac{\partial Q_j}{\partial H\mathit{\Lambda}_H}\frac{H\mathit{\Lambda}_H}{Q_j}\\
=&\frac{\partial A_jN_j^{\frac{\sigma}{\sigma-1}}}{\partial H\mathit{\Lambda}_H}\frac{H\mathit{\Lambda}_H}{Q_j}\\
=&A_jN_j^{\frac{\sigma}{\sigma-1}-1}\left(\theta_jH_j^{\frac{\sigma-1}{\sigma}-1}\frac{\partial H_j}{\partial H\mathit{\Lambda}_H}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}-1}\frac{\partial L_j}{\partial H\mathit{\Lambda}_H}\right)\frac{H\mathit{\Lambda}_H}{Q_j}\\
=&N_j^{\frac{\sigma}{\sigma-1}-2}\left(\theta_jH_j^{\frac{\sigma-1}{\sigma}}\frac{\partial\ln H_j}{\partial\ln H\mathit{\Lambda}_H}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}\right)\\
=&N_j^{\frac{\sigma}{\sigma-1}-2}\left(\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}+\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)\right)+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}\right)\\
=&N_j^{\frac{\sigma}{\sigma-1}-2}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)+N_j^{\frac{\sigma}{\sigma-1}-2}\left(\theta_jH_j^{\frac{\sigma-1}{\sigma}}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right)\frac{\partial\ln L_j}{\partial\ln H\mathit{\Lambda}_H}\\
=&N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)+\beta_L\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)\end{aligned}$$

This implies that

$$N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)+\beta_L\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}\right)=\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}$$

$$N_j^{\frac{\sigma}{\sigma-1}-1}N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)=\left(\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)(\beta_H-\beta_L)+\beta_L\right)\right)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}$$

$$N_j^{\frac{\sigma}{\sigma-1}-1}N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)=\left(\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_jH_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma\left(\beta_H-\beta_L\right)}{\sigma+\beta_H}\right)+\beta_L\right)\right)\frac{\partial\ln w_{Lj}}{\partial\ln H\mathit{\Lambda}_H}$$

$$\frac{N_j^{\frac{\sigma}{\sigma-1}-1}N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)}{\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)+\beta_L\right)}=\frac{\partial\ln w_{Lj}}{\partial\ln H\Lambda_H}$$

This implies that

$$\begin{aligned}
1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\Lambda_H} &= 1+\frac{N_j^{\frac{\sigma}{\sigma-1}-1}N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)(\beta_H-\beta_L)}{\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)+\beta_L\right)}\\
&= \frac{\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-\beta_L N_j^{\frac{\sigma}{\sigma-1}-1}}{\frac{1+\frac{\beta_L}{\sigma}}{\frac{1}{\sigma}-\frac{1}{\rho}}-N_j^{\frac{\sigma}{\sigma-1}-1}\beta_L-N_j^{\frac{\sigma}{\sigma-1}-1}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)\right)}\\
&= \frac{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}-\beta_L N_j^{\frac{1}{\sigma-1}}}{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}-N_j^{\frac{1}{\sigma-1}}\beta_L-N_j^{\frac{1}{\sigma-1}}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)\right)}\\
&= \frac{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}N_j^{\frac{1}{1-\sigma}}-\beta_L}{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}N_j^{\frac{1}{1-\sigma}}-\beta_L-\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)\right)}\\
&= \frac{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}A_j^{-\frac{1}{\sigma}}Q_j^{\frac{1}{\sigma}}-\beta_L}{\frac{\sigma+\beta_L}{1-\frac{\sigma}{\rho}}A_j^{-\frac{1}{\sigma}}Q_j^{\frac{1}{\sigma}}-\beta_L-\left(\frac{\theta_j H_j^{\frac{\sigma-1}{\sigma}}}{\theta_j H_j^{\frac{\sigma-1}{\sigma}}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}}\left(\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}\right)\right)}
\end{aligned}$$

$\square$

Proposition 2 highlights that whenever $\beta_H \neq \beta_L$, changes in wage index (Λ_H), labor supply of the high skilled H, and a_H have an opposite effect on the skill premium and skill ratio.[55] The statement also highlights that whenever the elasticity of substitution in production is roughly similar to the substitution elasticity across different type of goods $\sigma \approx \rho$ then $1+(\beta_H-\beta_L)\frac{\partial\ln w_{Lj}}{\partial\ln H\Lambda_H}\approx 1$ and so the effect of $\ln\Lambda_H H$ on skill ratio and skill premium is similar to Equations (3c) and (3d) in the main paper. Nevertheless, when σ and ρ are very different then the impact of $\ln\Lambda_H H$ on skill ratio and skill premium can potentially depend on firm-level characteristics such as $A_j^{-\frac{1}{\sigma}}Q_j^{\frac{1}{\sigma}}$ and $\frac{\theta_j H_j^{\frac{\sigma-1}{\sigma}}}{\theta_j H_j^{\frac{\sigma-1}{\sigma}}+(1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}}$. To deal with this issue empirically, we also explore robustness on whether interacting region-year fixed effects with pre-innovation share of high-skilled workers dummies substantially changes our results (see Table 3 in the main paper).

[55] It is easy to show that an analogous statement holds for Λ_L, L, and a_L.

C.3 Extension: Derivations with Capital in the Production Function

So far we have abstracted away from other inputs in the production function. Nevertheless, it is straightforward to extend the problem with other inputs. Here we demonstrate this by adding capital to the production function.

The new profit maximization problem is the following:

$$\pi_j(A_j, \theta_j) = \max_{w_{Hj}, w_{Lj}} p_j Q_j - H_j(w_{Hj})w_{Hj} - L_j(w_{Lj})w_{Lj} - rK_j \tag{C.19}$$

Subject to

$$Q_j = A_j \left(\left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}\frac{\varrho-1}{\varrho}} + K_j^{\frac{\varrho-1}{\varrho}} \right)^{\frac{\varrho}{\varrho-1}} \tag{C.20}$$

and budget constraints (C.4), (C.13), and (C.14).

The FOCs of the problem now become:

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\varrho-1}{\varrho}} Q_j^{\frac{1}{\varrho}} \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}} (1-\theta) L_j^{-\frac{1}{\sigma}} = \frac{1+\beta_L}{\beta_L} w_{Lj}$$

$$\frac{1+\rho}{\rho} p_j A_j^{\frac{\varrho-1}{\varrho}} Q_j^{\frac{1}{\varrho}} \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}} \theta H_j^{-\frac{1}{\sigma}} = \frac{1+\beta_H}{\beta_H} w_{Hj}$$

As a result, the ratio is the same as before

$$\frac{\theta_j H_j^{-\frac{1}{\sigma}} \frac{\beta_H}{1+\beta_H}}{(1-\theta_j) L_j^{-\frac{1}{\sigma}} \frac{\beta_L}{1+\beta_L}} = \frac{w_{Hj}}{w_{Lj}}$$

and so we get the same relationship between skill premium and skill demand as before (see Equation (C.11)):

$$\underbrace{\ln \frac{w_{Hj}}{w_{Lj}}}_{\substack{\text{skill} \\ \text{premium}}} = \underbrace{\ln \frac{1+\frac{1}{\beta_L}}{1+\frac{1}{\beta_H}}}_{\substack{\text{relative} \\ \text{mark-down}}} + \underbrace{\ln \frac{\theta_j}{1-\theta_j}}_{\substack{\text{skill} \\ \text{bias}}} - \underbrace{\frac{1}{\sigma} \ln \frac{H_j}{L_j}}_{\substack{\text{skill} \\ \text{ratio}}} \tag{C.21}$$

Note that Proposition 1 only uses this equation and Equations (C.13), and (C.14). And so the proposition can be proved by applying exactly the same steps.

Turning to Proposition 2, the first part of the statement says that changes in $X = \{H\Lambda_H, a_{Hj}\}$ have the following effect on the skill ratio and skill premium:

$$\frac{\partial \ln \frac{w_{Hj}}{w_{Lj}}}{\partial \ln X} = -\left(\frac{1}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln X}\right)$$

$$\frac{\partial \ln \frac{H_j}{L_j}}{\partial \ln X} = \left(\frac{\sigma}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln X}\right)$$

As for Proposition 1, this part of the statement only uses Equations (C.21), (C.13), and (C.14), which are unaffected by the presence of capital.

The proposition also derives $1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln X}$. The presence of capital changes the derivation of that part, which we develop here. The FOC for low-skilled workers in the presence of capital becomes

$$\ln\frac{1+\rho}{\rho} + \ln\left(\frac{I\kappa_j}{p^{1-\rho}}\right)^{\frac{1}{\rho}} + \frac{\varrho-1}{\varrho}\ln A_j + \left(\frac{1}{\varrho}-\frac{1}{\rho}\right)\ln Q_j + \frac{\varrho-\sigma}{\varrho(\sigma-1)}\ln\left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right] +$$

$$+\ln(1-\theta_j) - \frac{1}{\sigma}\ln L_j + \ln\frac{\beta_L}{1+\beta_L} = \ln w_{Lj}$$

And the FOC for capital is:

$$\ln\frac{1+\rho}{\rho} + \ln\left(\frac{I\kappa_j}{p^{1-\rho}}\right)^{\frac{1}{\rho}} + \frac{\varrho-1}{\varrho}\ln A_j + \left(\frac{1}{\varrho}-\frac{1}{\rho}\right)\ln Q_j - \frac{1}{\varrho}\ln K_j + \ln\left(1-\frac{1}{\rho}\right) = r \tag{C.22}$$

Take the derivative with respect to $\ln H\varLambda_H$, which leads to

$$\left(\frac{1}{\varrho}-\frac{1}{\rho}\right)\frac{\partial \ln Q_j}{\partial \ln H\varLambda_H} + \frac{\varrho-\sigma}{\varrho(\sigma-1)}\frac{\partial \ln\left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right]}{\partial \ln H\varLambda_H} - \frac{1}{\sigma}\frac{\partial \ln L_j}{\partial \ln H\varLambda_H} = \frac{\partial \ln w_{Lj}}{\partial \ln H\varLambda_H} \tag{C.23}$$

Now we want to express the three terms on the left hand side in the above equation in terms of $\frac{\partial \ln w_{Lj}}{\partial \ln H\varLambda_H}$. We denote $N_j = \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right]$ for notational convenience as before. For the first term we have:

$$\frac{\partial \ln Q_j}{\partial \ln H\varLambda_H} = \frac{\frac{\varrho}{\varrho-1}A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}}{1-A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}K_j^{\frac{\varrho-1}{\varrho}}\left(1-\frac{\varrho}{\rho}\right)}\cdot$$

$$\cdot\left(\frac{\varrho-1}{\varrho}N_j^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}+1}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\varLambda_H}\right)+\beta_L\frac{\partial \ln w_{Lj}}{\partial \ln H\varLambda_H}\right)\right)$$

where we used that Equation (C.22). This implies that $\left(1-\frac{\varrho}{\rho}\right)\frac{\partial \ln Q_j}{\partial \ln H\Lambda_H} = \frac{\partial \ln K_j}{\partial \ln H\Lambda_H}$.

For the second term in Equation (C.23), we obtain:

$$\frac{\partial \ln \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}}\right]}{\partial \ln H\Lambda_H} = \frac{\sigma-1}{\sigma}\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\left(1-\frac{\beta_H}{\sigma+\beta_H}\right)\left(1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right)+\beta_L\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H}\right)$$

Using as before that $\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H} = \frac{1}{\beta_L}\frac{\partial \ln L_j}{\partial \ln H\Lambda_H}$ for the third term, and plugging the three terms back into Equation (C.23) we get the following expression:

$$\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H} = \frac{\left(\frac{\left(\frac{1}{\varrho}-\frac{1}{\rho}\right)A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}N_j^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}+1}}{1-A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}K_j^{\frac{\varrho-1}{\varrho}}\left(1-\frac{\varrho}{\rho}\right)}+\frac{\varrho-\sigma}{\varrho\sigma}\right)N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\frac{\sigma}{\sigma+\beta_H}}{1+\beta_L-\left(\frac{\left(\frac{1}{\varrho}-\frac{1}{\rho}\right)A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}N_j^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}+1}}{1-A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}K_j^{\frac{\varrho-1}{\varrho}}\left(1-\frac{\varrho}{\rho}\right)}+\frac{\varrho-\sigma}{\varrho\sigma}\right)\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}+\beta_L\right)}$$

This implies that:

$$1+(\beta_H-\beta_L)\frac{\partial \ln w_{Lj}}{\partial \ln H\Lambda_H} = \frac{1+\beta_L-\left(\frac{\left(\frac{1}{\varrho}-\frac{1}{\rho}\right)A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}N_j^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}+1}}{1-A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}K_j^{\frac{\varrho-1}{\varrho}}\left(1-\frac{\varrho}{\rho}\right)}+\frac{\varrho-\sigma}{\varrho\sigma}\right)\beta_L}{1+\beta_L-\left(\frac{\left(\frac{1}{\varrho}-\frac{1}{\rho}\right)A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}N_j^{\frac{\varrho-\sigma}{\varrho(\sigma-1)}+1}}{1-A_j^{\frac{\varrho-1}{\varrho}}Q_j^{\frac{1-\varrho}{\varrho}}K_j^{\frac{\varrho-1}{\varrho}}\left(1-\frac{\varrho}{\rho}\right)}+\frac{\varrho-\sigma}{\varrho\sigma}\right)\left(N_j^{-1}\theta_j H_j^{\frac{\sigma-1}{\sigma}}\frac{\sigma(\beta_H-\beta_L)}{\sigma+\beta_H}+\beta_L\right)}$$

This expression is similar to the one that we obtained without capital in Proposition 2.

C.4 The derivation of the downward sloping firm-level demand function

We assume consumers in the market have preferences for love of variety

$$max_{\{Q_1,\ldots,Q_J\}}\left(\sum_j \kappa_j^{\frac{1}{\rho}} Q_j^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$$

subject to the following constraint:

$$\sum_j p_j Q_j = I$$

The Lagrangian of the problem is the following:

$$\mathcal{L} = \left(\sum_j \kappa_j^{\frac{1}{\rho}} Q_j^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} - \lambda\left(\sum_j p_j Q_j - I\right)$$

The FOC of this problem is the following:

$$\left(\sum_j \kappa_j^{\frac{1}{\rho}} Q_j^{\frac{\rho-1}{\rho}}\right)^{\frac{1}{\rho-1}} \kappa_j^{\frac{1}{\rho}} Q_j^{-\frac{1}{\rho}} - \lambda p_j = 0$$

and so

$$\left(\frac{\kappa_j}{\kappa_k}\right)^{\frac{1}{\rho}} \left(\frac{Q_j}{Q_k}\right)^{-\frac{1}{\rho}} = \frac{p_j}{p_k}$$

which can be rearranged to

$$Q_j = \frac{\kappa_j}{\kappa_k}\left(\frac{p_j}{p_k}\right)^{-\rho} Q_k$$

and

$$p_j Q_j = \frac{Q_k}{\kappa_k} p_k^{\rho} \kappa_j p_j^{1-\rho}$$

summing up this by j leads to the following equation

$$I = \sum_j p_j Q_j = \frac{Q_k}{\kappa_k} p_k^{\rho} \sum_j \kappa_j p_j^{1-\rho}$$

Let us define the price index as $p \equiv \left(\sum_j \kappa_j p_j^{1-\rho}\right)^{\frac{1}{1-\rho}}$ and then the above equation can be rewritten to

$$Q_j = \frac{I}{p^{1-\rho}} \kappa_j p_j^{-\rho}$$

which leads to the following demand equation for firm j:

$$\ln Q_j = \ln I - (1-\rho)\ln p + \ln \kappa_j - \rho \ln p_j$$

or

$$\ln p_j = \frac{1}{\rho}\ln I - \frac{1-\rho}{\rho}\ln p + \frac{1}{\rho}\ln \kappa_j - \frac{1}{\rho}\ln Q_j$$

or

$$p_j = \left(\frac{I\kappa_j}{p^{1-\rho}}\right)^{\frac{1}{\rho}} Q_j^{-\frac{1}{\rho}}.$$

Appendix D Extension: Labor Market Power

Throughout the paper (and in Appendix C) we have assumed that firms are atomistic and so they do not take into account that their actions potentially affect other firms' behavior. Deb et al. (2020) derive the impact of firm-level technological changes on relative wages and employment by taking into account strategic interactions between firms. This relationship is characterized by

$$\ln \frac{w_{H_{jt}}}{w_{L_{jt}}} = \ln \frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}} + \ln \frac{\theta_{jt}}{1-\theta_{jt}} - \frac{1}{\sigma} \ln \frac{H_{jt}}{L_{jt}}. \tag{D.1}$$

where $\frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$ captures the contribution of relative market power differences on skill premia. Deb et al. (2020) derive that in their model the effect of market power on wages can be expressed as follows:

$$\varepsilon_{Sjmt} = \frac{1}{\hat{\beta}_S} e_{Sjmt} + \frac{1}{\hat{\eta}_S}(1 - e_{Sjmt}) \tag{D.2}$$

where e_{Sjmt} is the market share of firm j of workers in skill group S in market m at time t;[56] while $\hat{\beta}_S$ and $\hat{\eta}_S$ are preference parameters of the consumers that determine the firm- and labor market-level labor supply elasticity in skill group S.[57] Notice that Equation (D.1) is very similar to Equation (C.11), derived in Appendix C focusing on the atomistic agents except that the relative mark-down term $\frac{1+\frac{1}{\beta_L}}{1+\frac{1}{\beta_H}}$ is now replaced with $\frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$. Crucially, in the atomistic case the relative mark-down $\frac{1+\frac{1}{\beta_L}}{1+\frac{1}{\beta_H}}$ is not firm-specific, but when we introduce strategic interactions, the relative markdown becomes firm-specific and depends on the firm's market share.

Following technological change or innovation, the change in Equation (D.1) will be

$$\underbrace{\Delta \ln \frac{w_{H_{jt}}}{w_{L_{jt}}}}_{\substack{\text{Change in} \\ \text{skill premium}}} = \underbrace{\Delta \ln \frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}}_{\substack{\text{Change in} \\ \text{markdown}}} + \underbrace{\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}}_{\substack{\text{Change in} \\ \text{skill bias}}} - \frac{1}{\sigma} \underbrace{\Delta \ln \frac{H_{jt}}{L_{jt}}}_{\substack{\text{Change in} \\ \text{skill ratio}}} \tag{D.3}$$

This equation is very similar to our benchmark equation (see Equation (5)) except for the extra

[56] Berger et al. (2019b) suggest to use the wage bill shares to calculate e_{Sjmt} when there are no productivity differences among workers. Nevertheless, if the productivity differences are large, the wage bill shares might simply be driven by those differences. As a result, we calculate market share based on market shares in terms of workers. Our results are robust to using the wage bill for calculating the market shares.

[57] Deb et al. (2020) present a model where $\hat{\beta}_S$ and $\hat{\mu}_S$ are the key parameters of the representative agent's labor supply function. Berger et al. (2019a) show in Appendix B.1 that such a representative agent's labor supply function can be micro-founded in a discrete choice framework as presented in our Section 2 and in Appendix C. When there are M distinct labor markets, the idiosyncratic preferences for working at a particular firm have the following type-I Extreme value distribution (where we applied our notation):

$$F(\epsilon_S ij, ..., \epsilon_S iJ) = \exp\left[- \sum_{m=1}^{M} \left(\sum_{j \in Market_m} e^{-(1+\hat{\beta_S})\epsilon_S ij} \right)^{\frac{1+\hat{\beta_S}}{1+\hat{\eta_S}}} \right].$$

When $\hat{\eta_S} = 0$, the distribution is the same as the one used in Appendix C. Whenever $\hat{\eta_S} > 0$, there is an increased correlation of draws within a labor market (McFadden et al. 1977), which creates a differential labor supply elasticity for moving across firms within a labor market, and moving across firms in different labor markets.

term that reflects the change in markdown coming from changes in labor market power (or rent sharing as called by Deb et al. 2020) following innovation. The intuition for that term is the following. When firms innovate, they might grow, which could potentially change their employment share in a given labor market and so their market power on that market. If the increase in market share differs between the college and non-college labor markets (or if the within- and between-market elasticities are different for college and non-college workers), then relative changes in market power will have a direct effect on the skill premia.

We will quantify the change in market power following innovation in two steps. First, we estimate the firm-level change in market shares using regression Equation (10). Since the definition of the "markets" is crucial for this exercise, we will explore various definitions of labor markets. Second, we use the parameter values for $\hat{\beta}_S$ and $\hat{\eta}_S$ from Deb et al. (2020) and calculate the firm-specific relative markdown, $\frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$, using Equation (D.2).

Panel A of Table D.1 summarizes the parameter values that we use in Equation (D.2) for calculating firm-level markdowns. In Panel B and C we report average markdowns for college and non-college workers under alternative labor market definitions. In Panel B we consider a local district-one digit industry combination as a labor market. In Panel C we follow Berger et al. (2019a) and apply a narrow definition with a combination of a district and a three-digit NACE industry.[58] Both definitions lead to very similar markdown estimates. The average markdown for college workers is between 0.60 and 0.65 and for non-college ones it is between 0.74 and 0.78 in both countries. The markdown is larger for college workers as their firm-level labor supply is less elastic.

Table D.2 shows the changes in market share and markdown following innovation. In rows (1)-(5), we define labor markets as one-digit NACE industry within a district (same as Table D.1 Panel A). Under this broader definition of labor markets we find no indication for any significant change in markdowns. In rows (5)-(10) we use a narrow definition of markets, where the college and non-college markets are defined as a three-digit NACE industry within a district (Table D.1 Panel C, following Berger et al. 2019a). When we use this narrow market definition, we find that the college and non-college share increases by roughly the same magnitude. Nevertheless, given that firm-level labor supply of non-college workers is more elastic, the change in market shares translates into a larger change in the markdown. Intuitively, non-college workers have weak preferences between firms, and so wage competition on that labor market is fiercer. As a result, gaining market power in that market makes a bigger difference. Row (10) demonstrates that, as a result, there is a 0.7% (s.e. 0.2%) increase in relative markdown in Norway if we apply this narrow definition of labor market. For Hungary, we find a 1.3% (s.e. 0.8%) increase in relative markdowns, which is only borderline significant.

This analysis highlights that relative changes in labor market power can only explain at most a small fraction of the change in skill premium observed in the data. In our preferred specification we estimate that the skill premium increased by 4.5% for Norway (see Column 4 of Table 2). This implies that at most 16% of the skill premium increase can be attributed to changes in market power. As a result, even if we incorporate the changes in firm-level markdowns into the calculation of firm-level

[58] In Norway we have 47 districts, while in Hungary we have 174. These are substantially smaller regional areas than commuting zones in the United States used by Berger et al. (2019a). As a result, our labor market definition is in fact narrower than the one used in Berger et al. (2019a).

changes in skill bias (see Section 5.4 and Equation (13)) we get very similar numbers.

Table D.1: Labor Market Power: Parameter Values and Descriptive Statistics

Variable	Value (NO)	Value (HU)	Description
Panel A: Parameter values			
$\hat{\eta_H}$	0.66	0.66	College workers' market-level labor supply elasticity
$\hat{\eta_L}$	0.66	0.66	Non-college workers' market-level labor supply elasticity
$\hat{\beta_H}$	1.85	1.85	College workers' firm-level labor supply elasticity
$\hat{\beta_L}$	8.12	8.12	Non-college workers' firm-level labor supply elasticity
Panel B: Average markdown (district × 1-digit industry)			
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1}{1+\varepsilon_{Hjmt}}$	0.64	0.63	Average markdown for college workers
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1}{1+\varepsilon_{Ljmt}}$	0.87	0.86	Average markdown for non-college workers
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$	0.74	0.74	Average relative markdown
Panel C: Average markdown (district × 3-digit industry)			
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1}{1+\varepsilon_{Hjmt}}$	0.60	0.53	Average markdown for college workers
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1}{1+\varepsilon_{Ljmt}}$	0.77	0.62	Average markdown for non-college workers
$\frac{1}{N_j}\sum_{j=1}^{N_j}\frac{1+\varepsilon_{Ljmt}}{1+\varepsilon_{Hjmt}}$	0.78	0.76	Average relative markdown

Notes: The parameter values come from Deb et al. (2020) who use between-market labor supply elasticites from Berger et al. (2019a). The labor market shares are calculated based on all firms in the employer-employee register for Norway and based on all firms in the Structure of Earnings Survey in Hungary. The average markdown is calculated for the firms in the CIS.

Table D.2: Change in Labor Market Power Following Firm-level Technological Change

Panel A: Norway

Measure	Level	Innovation	s.e.	Obs.	R-squared
(1) College market share	(CZ x 1-nace)	0.001	(0.001)	24,592	0.04
(2) Non-college market share	(CZ x 1-nace)	0.001*	(0.001)	24,959	0.04
(3) Log college markdown	(CZ x 1-nace)	-0.001	(0.001)	24,952	0.04
(4) Log non-college markdown	(CZ x 1-nace)	-0.001*	(0.001)	24,959	0.04
(5) Log relative markdown	(CZ x 1-nace)	0.001	(0.001)	24,952	0.08
(6) College market share	(CZ x 3-nace)	0.011***	(0.004)	24,301	0.11
(7) Non-college market share	(CZ x 3-nace)	0.012***	(0.004)	24,928	0.11
(8) Log college markdown	(CZ x 3-nace)	-0.005***	(0.002)	24,301	0.11
(9) Log non-college markdown	(CZ x 3-nace)	-0.010***	(0.003)	24,928	0.10
(10) Log relative markdown	(CZ x 3-nace)	0.007***	(0.002)	24,271	0.13

Panel B: Hungary

Measure	Level	Innovation	s.e.	Obs.	R-squared
(1) College market share	(CZ x 1-nace)	-0.003	(0.003)	2,357	0.28
(2) Non-college market share	(CZ x 1-nace)	-0.006**	(0.002)	2,364	0.32
(3) Log college markdown	(CZ x 1-nace)	0.001	(0.002)	2,357	0.26
(4) Log non-college markdown	(CZ x 1-nace)	0.005**	(0.002)	2.364	0.29
(5) Log relative markdown	(CZ x 1-nace)	0.005**	(0.002)	2,357	0.17
(6) College market share	(CZ x 3-nace)	0.014	(0.018)	1,995	0.12
(7) Non-college market share	(CZ x 3-nace)	-0.015	(0.013)	2,359	0.16
(8) Log college markdown	(CZ x 3-nace)	-0.006	(0.009)	1,995	0.11
(9) Log non-college markdown	(CZ x 3-nace)	0.015	(0.010)	2,359	0.15
(10) Log relative markdown	(CZ x 3-nace)	0.013*	(0.008)	1,990	0.13

Notes: This table shows the relationship between firm-level technological change and subsequent 6-year change in firms' market power. We measure firm-level technological change in the CIS survey which asks whether any new or significantly modified product/service/process/organizational change (aka innovation) was introduced. In the table each row reports the coefficients from regression Equation (10), where the dependent variable is in the first column of each row. In rows (1)-(5) the labor markets are defined at the district and 1-digit NACE industry level, while in rows (6)-(10) at the district and 3-digit NACE industry level. Relative markdowns are calculated based on Equation (D.2). In each regression we include log capital stock, log value added, the lagged dependent variable preceding the baseline year and industry-year fixed effects. Standard errors are clustered at the firm level and are reported in parentheses. Significance levels are: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$..

Appendix E Extension: Skill-biased technological change with bargaining

This section derives the relationship between skill ratio and skill premium by applying an alternative wage setting procedure. In the benchmark analysis we assumed wages determined based on the imperfect competition model proposed by Card et al. (2018). In Appendix D we presented an extension of the model where we allowed for strategic interaction between workers. Here we apply the bargaining model of Van Reenen (1996) and derive the optimal skill demand in that framework.

E.1 Wage Setting through Bargaining

Wage and employment determination.

Unions. We model wage and employment determination as a bargaining process between a firm and skill-specific unions. Assume that the union of workers with skill S at firm j has the following utility function (see Equation (1) in Van Reenen 1996):

$$U_{Sj} = S_j u(w_{Sj}) = S_j \frac{1}{1-m_S} w_{Sj}^{1-m_S} \tag{E.1}$$

where $0 \leq m_S \leq 1$ measures risk aversion of the workers that can vary by skill group S. This formulation reflect that unions care about not just the level of wages, but also about employment.

Firms. Firms' profit is given by the following function:

$$\Pi_j(A_j, \theta_j) = \max_{w_{Hj}, w_{Lj}} pQ_j - H_j w_{Hj} - L_j w_{Lj}$$

Subject to

$$\begin{aligned} Q_j &= A_j \left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j) L_j^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \\ \ln p_j &= \tfrac{1}{\rho}\ln \kappa_j - \tfrac{1}{\rho}\ln Q_j + \tfrac{\rho-1}{\rho}\ln p + \tfrac{1}{\rho}\ln I \end{aligned}$$

Similarly to Van Reenen (1996), we assume that firms are price takers on the output market. Nevertheless, it is straightforward to incorporate firms' price setting power into the framework presented here.

Equilibrium wage and employment. Wages are determined through a Nash-bargaining process. The equilibrium solution maximizes Ω by optimally choosing the skill-specific wages (w_{Hj} and w_{Lj}) and the skill specific employment (L_j and H_j) (see Equation (3) in Van Reenen 1996):

$$\max_{w_{Lj},w_{Hj},L_j,H_j,} \Omega = U_{Lj}^{\beta_L} U_{Hj}^{\beta_H} \Pi_j^{1-\beta_L-\beta_H} \tag{E.2}$$

where β_L and β_H are the bargaining powers of the two unions.

Solution. Plugging in U_{Lj} and U_{Hj} into the expression for Ω leads to the following formula:

$$\begin{aligned}\Omega &= \left[\tfrac{1}{1-m_L}\left(w_{Lj}\right)^{1-m_L} L_j\right]^{\beta_L} \times \left[\tfrac{1}{1-m_H}\left(w_{Hj}\right)^{1-m_H} H_j\right]^{\beta_H} \times \\ &\times \left[A_j\left[\theta_j H_j^{\frac{\sigma-1}{\sigma}} + (1-\theta_j)L_j^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} - H_j w_{Hj} - L_j w_{Lj}\right]^{1-\beta_L-\beta_H}\end{aligned}$$

The FOCs of this problem are the following:

$$\begin{aligned}
\frac{\partial\Omega}{\partial w_{Lj}} &= \Omega \times \left[\frac{\beta_L (w_{Lj})^{-m_L} L_j}{U_{Lj}} + \frac{-(1-\beta_L-\beta_H)L_j}{\Pi_j}\right] &= 0\\
\frac{\partial\Omega}{\partial w_{Hj}} &= \Omega \times \left[\frac{\beta_H (w_{Hj})^{-m_H} H_j}{U_{Hj}} + \frac{-(1-\beta_L-\beta_H)H_j}{\Pi_j}\right] &= 0\\
\frac{\partial\Omega}{\partial L_j} &= \Omega \times \left[\frac{\beta_L \frac{1}{1-m_L}(w_{Lj})^{1-m_L}}{U_{Lj}} + (1-\beta_L-\beta_H)\frac{PA_j^{\sigma}Q_j^{\frac{1}{\sigma}}(1-\theta_j)L_j^{-\frac{1}{\sigma}} - w_{Lj}}{\Pi_j}\right] &= 0\\
\frac{\partial\Omega}{\partial H_j} &= \Omega \times \left[\frac{\beta_H \frac{1}{1-m_H}(w_{Hj})^{1-m_H}}{U_{Hj}} + (1-\beta_L-\beta_H)\frac{PA_j^{\sigma}Q_j^{\frac{1}{\sigma}}\theta_j H_j^{-\frac{1}{\sigma}} - w_{Hj}}{\Pi_j}\right] &= 0
\end{aligned}$$

Rearranging and dividing the first and third and the second and fourth FOCs, we get:

$$\begin{aligned}
w_{Lj} &= \tfrac{1-m_L}{2-m_L} \times PA_j^{\sigma}Q_j^{\frac{1}{\sigma}}(1-\theta_j)L_j^{-\frac{1}{\sigma}}\\
w_{Hj} &= \tfrac{1-m_H}{2-m_L} \times PA_j^{\sigma}Q_j^{\frac{1}{\sigma}}\theta_j H_j^{-\frac{1}{\sigma}}
\end{aligned}$$

These equations show that both types of workers receive a share of their marginal product, which depends on their risk aversion parameter. Under risk neutrality, the marginal product is shared equally.

Dividing these two equations yields:

$$\frac{w_{Hj}}{w_{Lj}} = \frac{\frac{1-m_H}{2-m_H}}{\frac{1-m_L}{2-m_L}} \times \frac{\theta_j H_j^{-\frac{1}{\sigma}}}{(1-\theta_j)L_j^{-\frac{1}{\sigma}}} \tag{E.4}$$

The logarithm of this equation is the following

$$\ln\frac{w_{Hj}}{w_{Lj}} = \ln\frac{\frac{1-m_H}{2-m_H}}{\frac{1-m_L}{2-m_L}} + \ln\frac{\theta_j}{1-\theta_j} - \frac{1}{\sigma}\frac{H_j}{L_j} \tag{E.5}$$

The relative wage of the two types of workers depends on the relative marginal product and a wedge introduced by the bargaining process, when the risk aversion of the two types of workers is different. Since the wedge, $\frac{\frac{1-m_H}{2+m_H}}{\frac{1-m_L}{2+m_L}}$, depends only on the preference parameters of the workers (risk aversion of the high and low skilled workers), firm-level change in skill demand will have no effect on it.

Therefore, the change in the skill presmium following innovation will take the following form:

$$\underbrace{\Delta \ln \frac{w_{Hj}}{w_{Lj}}}_{\substack{\text{Change in} \\ \text{skill premium}}} = \underbrace{\Delta \ln \frac{\theta_j}{1-\theta_j}}_{\substack{\text{Change in} \\ \text{skill bias}}} - \underbrace{\frac{1}{\sigma} \Delta \ln \frac{H_j}{L_j}}_{\substack{\text{Change in} \\ \text{skill ratio}}} \tag{E.6}$$

This highlights that the relationship between change in skill premium and skill demand will be very similar in a bargaining model of wages and employment.

Appendix F Quantifying the Contribution of Firm-level Technological Change to the Economy-wide Skill Premium

This section studies the contribution of technological change to inequality. We assess the change in the wage premium between two periods, denoted by t and $t+1$, caused by firm-level technological change. Throughout the section we consider a hypothetical scenario: how the economy-wide wage premium would have evolved if there were no aggregate shift in skill supply (e.g. $H_t = H_{t+1}$ and $L_t = L_{t+1}$) or in quality of the workforce.

Our goal is to derive how the economy-wide skill premium is linked to firm-level technological change. We assume that wages have the following (log additive) structure:

$$\ln w_{it} = \alpha_t + \psi_i + \ln w_{Sj(i,t)} + \varepsilon_{it} \tag{F.1}$$

where i denotes workers and j denotes firms, ε_{it} is a mean zero error term. The ψ_i captures workers' skills that are portable across firms and are not affected by firm-level technological change (at least in the short term). The term $\ln w_{Sj(i,t)}$ represents the skill-group (S) specific firm-level wage premium that firm j pays. That wage premium depends on the technology applied by the firm.

The economy-wide college and non-college wage premia in this framework are given by the following equations:

$$\overline{\ln w_{H_t}} \equiv \frac{1}{H_t}\sum_{i\in H} \ln w_{it} = \alpha_t + \frac{1}{H_t}\sum_{i\in H}\psi_i + \frac{1}{H_t}\sum_{i\in H}\ln w_{Hj(i,t)}$$

$$\overline{\ln w_{L_t}} \equiv \frac{1}{L_t}\sum_{i\in L} \ln w_{it} = \alpha_t + \frac{1}{L_t}\sum_{i\in L}\psi_i + \frac{1}{L_t}\sum_{i\in L}\ln w_{Lj(i,t)}$$

The contribution of technological change to the economy-wide wage premium comes from the changes in $\frac{1}{H_t}\sum_{i\in H}\ln w_{Hj(i,t)}$ and $\frac{1}{L_t}\sum_{i\in L}\ln w_{Lj(i,t)}$. We can express that in the following way:

$$
\begin{aligned}
\Delta\left(\overline{\ln w_{H_t}}-\overline{\ln w_{L_t}}\right) &= \frac{1}{H_{t+1}}\sum_{i\in H}\ln w_{Hj(i,t+1)}-\frac{1}{H_t}\sum_{i\in H}\ln w_{Hj(i,t)}- \\
&\quad -\left(\frac{1}{L_{t+1}}\sum_{i\in L}\ln w_{Lj(i,t+1)}-\frac{1}{L_t}\sum_{i\in L}\ln w_{Lj(i,t)}\right) \\
&= \sum_j \frac{H_{jt+1}}{H_{t+1}}\ln w_{Hjt+1}-\sum_j\frac{H_{jt}}{H_t}\ln w_{Hjt}- \\
&\quad -\left(\sum_j\frac{L_{jt+1}}{L_{t+1}}\ln w_{Ljt+1}-\sum_j\frac{L_{jt}}{L_t}\ln w_{Ljt}\right) \\
&= \sum_j\left(\frac{H_{jt+1}}{H_{t+1}}-\frac{H_{jt}}{H_t}\right)\ln w_{Hjt+1}+\sum_j\frac{H_{jt}}{H_t}\left(\ln w_{Hjt+1}-\ln w_{Hjt}\right)- \\
&\quad -\left(\sum_j\left(\frac{L_{jt+1}}{L_{t+1}}-\frac{L_{jt}}{L_t}\right)\ln w_{Ljt+1}+\sum_j\frac{L_{jt}}{L_t}\left(\ln w_{Ljt+1}-\ln w_{Ljt}\right)\right)
\end{aligned}
$$

This leads us to Equation (15) in the main text:

$$
\begin{aligned}
\Delta\Theta=\Delta\left(\overline{\ln w_{H_t}}-\overline{\ln w_{L_t}}\right) &= \underbrace{\sum_j\left(\frac{H_{jt+1}}{H_{t+1}}-\frac{H_{jt}}{H_t}\right)\ln w_{Hjt+1}-\sum_j\left(\frac{L_{jt+1}}{L_{t+1}}-\frac{L_{jt}}{L_t}\right)\ln w_{Ljt+1}}_{\text{Reallocation effect}}+ \\
&\quad \underbrace{+\sum_j\frac{H_{jt}}{H_t}\left(\ln w_{Hjt+1}-\ln w_{Hjt}\right)-\sum_j\frac{L_{jt}}{L_t}\left(\ln w_{Ljt+1}-\ln w_{Ljt}\right)}_{\text{Wage premium effect}}
\end{aligned}
\tag{F.2}
$$

Let us distinguish between two types of firms: firms that change their technology (aka innovate) between t and $t+1$, denoted by *inn*; and others (non-innovators), denoted by *non*. Let us define the (baseline) weighted average skill premium for skill group $S\in\{L,H\}$ at time $t+1$ for the innovative and non-innovative firms to be the following:

$$
\begin{aligned}
\overline{\ln w_{Sjt+1}}^{inn} &\equiv \frac{\sum_{j\in inn} S_{jt}\ln w_{Sjt+1}}{\sum_{j\in inn} S_{jt}} \\
\overline{\ln w_{Sjt+1}}^{non} &\equiv \frac{\sum_{j\in non} S_{jt}\ln w_{Sjt+1}}{\sum_{j\in non} S_{jt}}
\end{aligned}
$$

Let us analyze first the reallocation term in Equation (F.2). The change in shares for the two skill groups can be also rewritten as

$$\begin{aligned}
\frac{S_{jt+1}}{S_{t+1}} - \frac{S_{jt}}{S_t} &= \frac{S_{jt} + \frac{S_{jt+1}-S_{jt}}{S_{jt}} S_{jt}}{S_t + \frac{S_{t+1}-S_t}{S_t} S_t} - \frac{S_{jt}}{S_t} \\
&= \frac{S_{jt} + \frac{S_{t+1}-S_t}{S_t} S_{jt} + \left(\frac{S_{jt+1}-S_{jt}}{S_{jt}} - \frac{S_{t+1}-S_t}{S_t}\right) S_{jt}}{S_t + \frac{S_{t+1}-S_t}{S_t} S_t} - \frac{S_{jt}}{S_t} \\
&= \left(\frac{S_{jt+1}-S_{jt}}{S_{jt}} - \frac{S_{t+1}-S_t}{S_t}\right) \frac{S_{jt}}{S_{t+1}} \\
&= \Delta s_j \frac{S_{jt}}{S_{t+1}}
\end{aligned} \tag{F.3}$$

where $\Delta s_j = \frac{S_{jt+1}-S_{jt}}{S_{jt}} - \frac{S_{t+1}-S_t}{S_t}$ shows the percent change in the number of workers in skill group S in firm j relative to the aggregate change in the number of workers in that skill group. Similarly to the skill premium, we can also define the average (baseline) change at time $t+1$ for the innovative and non-innovative firms to be the following:

$$\begin{aligned}
\overline{\Delta s_j}^{inn} &\equiv \frac{\sum_{j \in inn} \Delta s_j}{J^{inn}} \\
\overline{\Delta s_j}^{non} &\equiv \frac{\sum_{j \in non} \Delta s_j}{J^{non}}
\end{aligned}$$

where J^{inn} and J^{non} refer to the total number of innovative and non-innovative firms, respectively.

Without loss of generality we further assume that the change in employment share, Δs_j, is unrelated to skill share, S_{jt}/S_{t+1}, and the skill premium, $\ln w_{Sjt+1}$, within the two firm types and so the following will hold:[59]

$$\begin{aligned}
\sum_j \left(\frac{S_{jt+1}}{S_{t+1}} - \frac{S_{jt}}{S_t}\right) \ln w_{Sjt+1} &= \sum_{j \in inn} \Delta s_j \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1} + \sum_{j \in non} \Delta s_j \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1} \\
&= \frac{\sum_{j \in inn} \Delta s_j}{J^{inn}} \times \sum_{j \in inn} \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1} + \\
&\quad + \frac{\sum_{j \in non} \Delta s_j}{J^{non}} \times \sum_{j \in non} \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1} \\
&= \frac{\sum_{j \in inn} \Delta s_j}{J^{inn}} \times \frac{\sum_{j \in inn} \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1}}{\sum_{j \in inn} \frac{S_{jt}}{S_{t+1}}} \sum_{j \in inn} \frac{S_{jt}}{S_{t+1}} + \\
&\quad + \frac{\sum_{j \in non} \Delta s_j}{J^{non}} \times \frac{\sum_{j \in non} \frac{S_{jt}}{S_{t+1}} \ln w_{Sjt+1}}{\sum_{j \in non} \frac{S_{jt}}{S_{t+1}}} \sum_{j \in non} \frac{S_{jt}}{S_{t+1}} \\
&= \overline{\Delta s_j}^{inn} \times \overline{\ln w_{Sjt+1}}^{inn} \times \vartheta_{Sjt}^{inn} + \overline{\Delta s_j}^{non} \times \overline{\ln w_{Sjt+1}}^{non} \times \vartheta_{Sjt}^{non}
\end{aligned} \tag{F.4}$$

where $\vartheta_{Sjt}^{inn} \equiv \sum_{j \in inn} \frac{S_{jt}}{S_{t+1}}$ and $\vartheta_{Sjt}^{non} \equiv \sum_{j \in non} \frac{S_{jt}}{S_{t+1}}$. The above formula highlights that reallocation effects for skill groups S will depend on the percent change in employment shares from skill S at

[59] If this did not hold, we would simply need to break the innovative and non-innovative firms into further subgroups until we get this assumption to hold. Then we simply need to calculate the change in employment for each relevant subgroups and the skill premium in those subgroups. While applying it to more than two groups of firms (e.g. innovative and non-innovate) involves more notation, the same result can be obtained. The reallocation effects will be the change in share for each relevant subgroups multiplied by the average wage premium paid in each group.

innovative and non innovative firms, the wage premium paid by innovative and non-innovative firms, and the initial share of innovative and non-innovative firms.

This formula can be further simplified if we consider the effect of reallocation in absence of any change in aggregate supply of skills – meaning that $S_t = S_{t+1}$. In that case we have the following relationship between $\overline{\Delta s_j}^{inn}$ and $\overline{\Delta s_j}^{non}$:

$$\begin{aligned}
\overline{\Delta s_j}^{inn} \times \vartheta_{Sjt}^{inn} &= \tfrac{1}{J^{inn}} \textstyle\sum_{j \in inn} \frac{S_{jt+1} - S_{jt}}{S_{jt}} \times \sum_{j \in inn} \frac{S_{jt}}{S_t} \\
&= \textstyle\sum_{j \in inn} \Delta s_j \frac{S_{jt}}{S_t} \\
&= \textstyle\sum_{j \in inn} \frac{S_{jt+1} - S_{jt}}{S_{jt}} \frac{S_{jt}}{S_t} \\
&= \textstyle\sum_{j \in inn} \frac{S_{jt+1} - S_{jt}}{S_t} \\
&= -\textstyle\sum_{j \in non} \frac{S_{jt+1} - S_{jt}}{S_t} \\
&= -\textstyle\sum_{j \in non} \frac{S_{jt+1} - S_{jt}}{S_{jt}} \frac{S_{jt}}{S_t} \\
&= -\textstyle\sum_{j \in non} \Delta s_j \frac{S_{jt}}{S_t} \\
&= -\tfrac{1}{J^{inn}} \textstyle\sum_{j \in non} \Delta s_j \times \sum_{j \in non} \frac{S_{jt}}{S_t} \\
&= -\overline{\Delta s_j}^{non} \times \vartheta_{Sjt}^{non}
\end{aligned} \tag{F.5}$$

As a result, each term in the reallocation effect in Equation (F.2) can be rewritten as

$$\sum_j \left(\frac{S_{jt+1}}{S_{t+1}} - \frac{S_{jt}}{S_t} \right) \ln w_{Sjt+1} = \underbrace{\overline{\Delta s_j}^{inn} \times \vartheta_{Sjt}^{inn}}_{\text{Change in share of inn firms}} \times \underbrace{\left(\overline{\ln w_{Sjt+1}}^{inn} - \overline{\ln w_{Sjt+1}}^{non} \right)}_{\text{Difference in wage premia between inn/non}} \tag{F.6}$$

Based on this derivation the reallocation effect in Equation (F.2) will be the following

$$\begin{aligned}
\text{Reallocation eff.} \quad = \quad & \underbrace{\overline{\Delta h_j}^{inn} \times \vartheta_{Hjt}^{inn}}_{\text{Change in H share of inn firms}} \times \underbrace{\left(\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non} \right)}_{\text{Difference in H wage premia between inn/non}} - \\
- \quad & \underbrace{\overline{\Delta l_j}^{inn} \times \vartheta_{Ljt}^{inn}}_{\text{Change in L share of inn firms}} \times \underbrace{\left(\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non} \right)}_{\text{Difference in L wage premia between inn/non}}
\end{aligned} \tag{F.7}$$

According to this equation, the reallocation effect depends on the market share of innovative firms in the two labor markets $\left(\vartheta_{Hjt}^{inn}, \vartheta_{Ljt}^{inn}\right)$, the proportional increase in the number of workers in innovative firms in the two markets $\left(\overline{\Delta h_j}^{inn}, \overline{\Delta l_j}^{inn}\right)$ and the premia innovative firms pay in the two markets. We will apply this formula to assess the contribution of technological change via the reallocation term to overall inequality.

Let us turn to the wage premium effect. As we derived in Appendix C, if firm j changes its technology, we will have the following change in the firm-level skill premium:

$$\ln \frac{w_{Hjt+1}}{w_{Ljt+1}} - \ln \frac{w_{Hjt}}{w_{Ljt}} = \ln \frac{\theta_{jt+1}}{1 - \theta_{jt+1}} - \ln \frac{\theta_{jt}}{1 - \theta_{jt}} - \frac{1}{\sigma} \left[\ln \frac{H_{jt+1}}{L_{jt+1}} - \ln \frac{H_{jt}}{L_{jt}} \right] \tag{F.8}$$

As a result, the change in the skill premium in innovative and non-innovative firms is, according to Equation (F.8):

$$\begin{array}{lcrcl} \text{Innovative} & : & \Delta \ln w_{Hjt} - \Delta \ln w_{Ljt} & = & \Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}} - \frac{1}{\sigma}\left(\Delta \ln H_{jt} - \Delta \ln L_{jt}\right) \\ \text{Non-innovative} & : & \Delta \ln w_{Hjt} - \Delta \ln w_{Ljt} & = & -\frac{1}{\sigma}\left(\Delta \ln H_{jt} - \Delta \ln L_{jt}\right) \end{array}$$

Substituting these into the formula of the wage premium effect:

$$\begin{aligned} \text{Wage premium effect} &= \textstyle\sum_j \frac{H_{jt}}{H_t}\left(\ln w_{Hjt+1} - \ln w_{Hjt}\right) - \sum_j \frac{L_{jt}}{L_t}\left(\ln w_{Ljt+1} - \ln w_{Ljt}\right) \\ &= \textstyle\sum_j \frac{H_{jt}}{H_t}\left(\Delta \ln w_{Hjt} - \Delta \ln w_{Ljt}\right) + \sum_j \left(\frac{H_{jt}}{H_t} - \frac{L_{jt}}{L_t}\right) \Delta \ln w_{Ljt} \\ &= \textstyle\sum_{j\in inn} \frac{H_{jt}}{H_t}\left(\Delta \ln w_{Hjt} - \Delta \ln w_{Ljt}\right) + \sum_{j\in non} \frac{H_{jt}}{H_t}\left(\Delta \ln w_{Hjt} - \Delta \ln w_{Ljt}\right) + \\ &\quad \textstyle + \sum_j \left(\frac{H_{jt}}{H_t} - \frac{L_{jt}}{L_t}\right) \Delta \ln w_{Ljt} \\ &= \textstyle\sum_{j\in inn} \frac{H_{jt}}{H_t}\left(\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}} - \frac{1}{\sigma}\left(\Delta \ln H_{jt} - \Delta \ln L_{jt}\right)\right) + \\ &\quad \textstyle + \sum_{j\in non} \frac{H_{jt}}{H_t}\left(-\frac{1}{\sigma}\left(\Delta \ln H_{jt} - \Delta \ln L_{jt}\right)\right) + \\ &\quad \textstyle + \sum_j \left(\frac{H_{jt}}{H_t} - \frac{L_{jt}}{L_t}\right) \Delta \ln w_{Ljt} \\ &= \underbrace{\sum_{j\in inn} \frac{H_{jt}}{H_t}\left[\Delta \ln \frac{\theta_{jt}}{1-\theta_{jt}}\right]}_{\text{Direct effect of skill bias}} - \\ &\quad \underbrace{-\frac{1}{\sigma}\left[\sum_j \frac{H_{jt}}{H_t}\left(\Delta \ln H_{jt} - \Delta \ln L_{jt}\right)\right]}_{\text{Average change in log skill ratio}} - \\ &\quad - \underbrace{\sum_j \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \Delta \ln w_{Ljt}}_{\substack{\text{Change in low skilled premium weighted by} \\ \text{the difference between high and low} \\ \text{skill employment share}}} \end{aligned} \tag{F.9}$$

The first term in this equation, the direct effect of skill bias, can be rewritten as

$$
\begin{aligned}
\sum_{j\in inn} \frac{H_{jt}}{H_t}\left[\Delta\ln\frac{\theta_{jt}}{1-\theta_{jt}}\right] &= \frac{\sum_{j\in inn} H_{jt}\left[\Delta\ln\frac{\theta_{jt}}{1-\theta_{jt}}\right]}{H_t} \\
&= \frac{\sum_{j\in inn} H_{jt}}{H_t}\frac{\sum_{j\in inn} H_{jt}\left[\Delta\ln\frac{\theta_{jt}}{1-\theta_{jt}}\right]}{\sum_{j\in inn} H_{jt}} \\
&= \frac{H_{t+1}}{H_t}\vartheta^{inn}_{Hjt}\times\Delta\ln\frac{\theta}{1-\theta} \\
&= \underbrace{\vartheta^{inn}_{Hjt}}_{\substack{\text{Share of}\\ \text{inn firms}}} \times \underbrace{\Delta\ln\frac{\theta}{1-\theta}}_{\substack{\text{Average change}\\ \text{in skill bias}}}
\end{aligned}
\tag{F.10}
$$

where the last equality takes into account that $H_{t+1} = H_t$. The (weighted) average change in skill bias, $\Delta\ln\frac{\theta}{1-\theta} \equiv \frac{\sum_{j\in inn} H_{jt}\left[\Delta\ln\frac{\theta_{jt}}{1-\theta_{jt}}\right]}{\sum_{j\in inn} H_{jt}}$, is defined in Equation (13) in Section 5.4.

If $H_{t+1} = H_t$ and $L_{t+1} = L_t$, the second term in Equation (F.9) can be written as

$$
\begin{aligned}
\sum_j \frac{H_{jt}}{H_t}(\Delta\ln H_{jt} - \Delta\ln L_{jt}) &\approx \sum_j \frac{H_{jt}}{H_t}\frac{H_{jt+1}-H_{jt}}{H_{jt}} - \sum_j \frac{H_{jt}}{H_t}\frac{L_{jt+1}-L_{jt}}{L_{jt}} \\
&= \sum_j \frac{H_{jt}}{H_t}\frac{H_{jt+1}-H_{jt}}{H_{jt}} - \sum_j \frac{L_{jt}}{L_t}\frac{L_{jt+1}-L_{jt}}{L_{jt}} - \sum_j\left(\frac{H_{jt}}{H_t}-\frac{L_{jt}}{L_t}\right)\frac{L_{jt+1}-L_{jt}}{L_{jt}} \\
&= \sum_j \frac{H_{jt+1}-H_{jt}}{H_t} - \sum_j \frac{L_{jt+1}-L_{jt}}{L_t} + \sum_j\left(\frac{L_{jt}}{L_t}-\frac{H_{jt}}{H_t}\right)\frac{L_{jt+1}-L_{jt}}{L_{jt}} \\
&= 0 - 0 + \sum_{j\in inn}\left(\frac{L_{jt}}{L_t}-\frac{H_{jt}}{H_t}\right)\frac{L_{jt+1}-L_{jt}}{L_{jt}} + \sum_{j\in non}\left(\frac{L_{jt}}{L_t}-\frac{H_{jt}}{H_t}\right)\frac{L_{jt+1}-L_{jt}}{L_{jt}} \\
&= \sum_{j\in inn}\left(\frac{L_{jt}}{L_t}-\frac{H_{jt}}{H_t}\right)\times\frac{1}{J_{inn}}\sum_{j\in inn}\frac{L_{jt+1}-L_{jt}}{L_{jt}} + \\
&\quad + \sum_{j\in non}\left(\frac{L_{jt}}{L_t}-\frac{H_{jt}}{H_t}\right)\times\frac{1}{J_{non}}\sum_{j\in non}\frac{L_{jt+1}-L_{jt}}{L_{jt}} \\
&= \left(\vartheta^{inn}_{Ljt}-\vartheta^{inn}_{Hjt}\right)\times\overline{\Delta l_j}^{inn} + \left(\vartheta^{non}_{Ljt}-\vartheta^{non}_{Hjt}\right)\times\overline{\Delta l_j}^{non} \\
&= \underbrace{\left(\vartheta^{inn}_{Ljt}-\vartheta^{inn}_{Hjt}\right)}_{\substack{\text{Difference between}\\ \text{inn firms' share}\\ \text{in H/L market}}} \times \underbrace{\left(\overline{\Delta l_j}^{inn}-\overline{\Delta l_j}^{non}\right)}_{\substack{\text{Diff between av.}\\ \text{growth rate in L}\\ \text{workers between inn./non}}}
\end{aligned}
\tag{F.11}
$$

where in the first approximation we have used that the log changes in skill S can be expressed as[60]

$$\Delta\ln S_{jt} = \ln S_{jt+1} - \ln S_{jt} \approx \frac{S_{jt+1}-S_{jt}}{S_{jt}}.$$

In the last but two equality in Equation (F.11) we assumed that among innovative and non-innovative firms, the change in low skilled employment is unrelated to the initial high skill share at those firms.

[60]The approximation comes from a first-order Taylor approximation showing that percentage and log percentage changes are similar when the change is small.

If this assumption does not hold, we need simply to disaggregate further until the assumption holds (see footnote 58 for further details). In the last equality we used the fact that $\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn} = -\left(\vartheta_{Ljt}^{non} - \vartheta_{Hjt}^{non}\right)$.

This result shows that the second term in the wage premium effect is the difference between the market share of innovative firms in the high- and low-skilled market multiplied by the difference in the average growth rate of low-skilled workers between innovative and non-innovative firms.

The third term in Equation (F.9) shows the correlation between the difference in firm's share in the high vs low-skilled market in t and the change in the low-skilled premia.

Without loss of generality, let us assume that, within the groups of innovative and non-innovative firms, the change in the wage of low-skilled workers, $\Delta \ln w_{Ljt}$, is independent from the number of low- and high-skilled workers in the firm. Again if this does not hold, we need to apply more subgroups of firms (see footnote 59). The formula for the third term can be rewritten as:

$$\begin{aligned}
\sum_j \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \Delta \ln w_{Ljt} &= \sum_{j\in inn} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \Delta \ln w_{Ljt} + \sum_{j\in non} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \Delta \ln w_{Ljt} \\
&= \sum_{j\in inn} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{inn}} \sum_{j\in inn} \Delta \ln w_{Ljt} + \\
&\quad + \sum_{j\in non} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt} \\
&= \sum_{j\in inn} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{inn}} \sum_{j\in inn} \Delta \ln w_{Ljt} - \\
&\quad - \sum_{j\in inn} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt} + \\
&\quad + \sum_{j\in inn} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt} + \\
&\quad + \sum_{j\in non} \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt} \\
&= \left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right) \times \left(\frac{1}{J_{inn}} \sum_{j\in inn} \Delta \ln w_{Ljt} - \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt}\right) + \\
&\quad + \sum_j \left(\frac{L_{jt}}{L_t} - \frac{H_{jt}}{H_t}\right) \times \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt} \\
&= \underbrace{\left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right)}_{\substack{\text{Difference between} \\ \text{inn firms' share} \\ \text{in H/L market}}} \times \underbrace{\left(\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}\right)}_{\substack{\text{Difference of av. L wage} \\ \text{changes between inn/non}}} + 0
\end{aligned}$$

where $\overline{\Delta w_{Lj}}^{inn} = \frac{1}{J_{inn}} \sum_{j\in inn} \Delta \ln w_{Ljt}$ and $\overline{\Delta w_{Lj}}^{non} = \frac{1}{J_{non}} \sum_{j\in non} \Delta \ln w_{Ljt}$ are the (unweighted) average growth rates of low-skilled wages in innovative and non-innovative firms, respectively. In the second equality we used that the low-skilled wage changes are independent of the initial number of high- and low-skilled workers within innovative and non-innovative firms.

These results imply that the wage premium effect will be given by the following equation whenever

$H_t = H_{t+1}$:

$$\text{Wage premium eff.} \quad = \quad \underbrace{\vartheta^{inn}_{Hjt} \times \Delta \ln \frac{\theta}{1-\theta}}_{\text{Direct effect of skill bias}} - \\ -\frac{1}{\sigma} \underbrace{\left(\vartheta^{inn}_{Ljt} - \vartheta^{inn}_{Hjt}\right) \times \left(\overline{\Delta l_j}^{inn} - \overline{\Delta l_j}^{non}\right)}_{\text{Average change in log skill ratio}} \\ - \underbrace{\left(\vartheta^{inn}_{Ljt} - \vartheta^{inn}_{Hjt}\right) \times \left(\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}\right)}_{\substack{\text{Change in low skilled premium weighted by} \\ \text{the difference between high and low} \\ \text{skill employment share}}} \tag{F.12}$$

These insights allow us to write up the effect of technological change on inequality as:

$$\Delta\Theta = \Delta\left(\overline{\ln w_{H_t}} - \overline{\ln w_{L_t}}\right) \quad =$$

$$\text{Reallocation eff.} \quad \left\{ \begin{array}{l} + \underbrace{\overline{\Delta h_j}^{inn} \times \vartheta^{inn}_{Hjt}}_{\text{Change in H share of inn firms}} \times \underbrace{\left(\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non}\right)}_{\text{Difference in H wage premia between inn/non}} \\ - \underbrace{\overline{\Delta l_j}^{inn} \times \vartheta^{inn}_{Ljt}}_{\text{Change in L share of inn firms}} \times \underbrace{\left(\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non}\right)}_{\text{Difference in L wage premia between inn/non}} \end{array} \right.$$

$$+ \text{ Wage premium eff.} \quad \left\{ \begin{array}{l} + \underbrace{\vartheta^{inn}_{Hjt} \times \Delta \ln \frac{\theta}{1-\theta}}_{\text{Direct effect of skill bias}} \\ - \frac{1}{\sigma} \underbrace{\left(\vartheta^{inn}_{Ljt} - \vartheta^{inn}_{Hjt}\right) \times \left(\overline{\Delta l_j}^{inn} - \overline{\Delta l_j}^{non}\right)}_{\text{Average change in log skill ratio}} \\ - \underbrace{\left(\vartheta^{inn}_{Ljt} - \vartheta^{inn}_{Hjt}\right) \times \left(\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}\right)}_{\substack{\text{Change in low skilled premium weighted by} \\ \text{the difference between high and low} \\ \text{skill employment share}}} \end{array} \right. \tag{F.13}$$

F.1 Empirical implementation

We use Equation (F.13) to quantify the extent to which firm-level technological change contributes to the aggregate college premium. Table F.1 summarizes how we calculate each of the components in Equation (F.13).

The $\overline{\Delta h_j}^{inn}$ and $\overline{\Delta l_j}^{inn}$ objects are just the proportional changes in skilled and unskilled workers in innovative firms, respectively. We calculate these from the firm-level regressions on employment growth and the change in skill ratio. $\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non}$ and $\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non}$, which are also part of the reallocation term, show the wage difference of college and non-college workers between innovative and non-innovative firms. Here we would like to filter out workers' composition effects—we are interested in how a particular worker's wage would change if she moved to an innovative firm. Therefore we start from column (2) of Table 2, but we also include worker fixed effects in Norway.

The shares of innovative firms in terms of college and non-college workers, ϑ_{Hjt}^{inn} and ϑ_{Ljt}^{inn}, are obtained from the CIS. We calculate the share of high-skilled workers at the innovative firms and apply the sampling weights provided with the CIS survey.[61]

The extent of skill-bias change, $\Delta \ln \frac{\theta}{1-\theta}$, is calculated based on Equation (13), where we include our preferred estimates for innovation's effect on the skill premium and the skilled share. We use these quantities together with the change in the number of unskilled workers in innovative firms to calculate the change in the number of these workers in non-innovative firms, $\overline{\Delta l_j}^{non}$.

Finally, $\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}$, the average wage increase of low-skilled workers in innovative firms relative to non-innovative firms, comes from the estimated coefficient of "Innovation" in our preferred specifications, Column (4) of Table 2.

The only component in (F.13) that we do not estimate in our data is the elasticity of substitution between high and low skilled workers, σ. For that variable we take the estimated values from the literature. Autor et al. (2003) argue that the elasticity between college and non-college workers is 2.94. If we apply that value of elasticity and calculate the contribution of firm-level technological change to economy-wide inequality, we can explain the change in college premium and college ratio observed in the data (see Section 6.1 for details). Furthermore, in Table F.4 we explore the sensitivity of our estimates for various values of σ. Reassuringly, the estimated magnitudes are not sensitive to the specific value of σ used.

Table F.2 shows the specific value of each component. Rows (1)-(2) show that the number of skilled workers employed by innovative firms increases substantially in both countries, while the number of unskilled workers in the same firms tend to decrease slightly. Rows (3)-(4) show the premia payed by innovative firms to high- and low-skilled workers. Both types of workers earn more in innovative firms, with the difference being substantially larger in Hungary.[62] Rows (5)-(6) show the share of firms conducting different types of innovation in the high skilled labor market. Row (7) shows the change in the number of non-college workers in non-innovative firms followed by the estimated skill bias in row (8). The final row shows our estimates for the difference in low-skilled workers' wage increase between innovative and non-innovative firms.

We use Equation (F.13) to calculate the contribution of technological change to the skill premium from these components. As both the shares and the coefficients reflect innovation activities conducted

[61] These weights are not available for Hungary, where we report unweighted results.

[62] Note that this is likely to be an overestimate in Hungary. This is because that, unlike in Norway, we cannot include worker fixed effects in our regression when estimating the premia innovative firms pay for college and non-college workers. As a more conservative approach, we compare the coefficients with and without worker fixed effects, and rescale the Hungarian premium by a similar factor. Including worker fixed effects reduces the estimated premium of non-college workers by 38% and the skill premium by 45%. Reducing the premia to a similar degree in Hungary reduces the reallocation effect to 2.33 pp.

over a 7-year period, these estimates also show the effect of innovation taking place over a 7-year period. For easier interpretation, we convert these to reflect a 10-year period.

Besides the overall contribution of all firm innovation to the increase in the college premium, we are also interested in the contribution of different forms of technological change. We split up firms along three lines: (i) whether they conduct R&D; (ii) whether their innovation is "new to market" and (iii) whether they conduct technical innovation or organizational change or both. We follow the same approach and estimate the skill bias separately for each group of innovative firms, similarly to Section A.6.

The results of the decomposition are presented in Table F.3. Let us start with column (1), which shows the overall effect of innovation. For both countries, the first row shows the skill bias. The next row is the reallocation effect, which contributed to the increase in skill premium by 0.52 and 3.74 pp. during a 10-year period in Norway and Hungary, respectively. The wage premium effect was 5.58 pp in Norway and 10.09 pp in Hungary. The total effect is just the sum of the reallocation and wage premium effects.

According to our results, skill-biased innovation contributed to the increase in the aggregate skill premium by 0.6 and 1.4 percentage points per year in Norway and Hungary, respectively. The bulk of the contribution results from the wage premium effect, suggesting that innovation mainly contributes to the aggregate skill premium via within-firm wage premia changes rather than the reallocation of workers to those firms. Within the wage premium effect, the direct effect dominates (Table F.3).[63] The higher contribution in Hungary is much in line with technology adoption generating more skill bias in Hungary compared to Norway, which is closer to the technological frontier.

We also show the contribution of different forms of technological change in Table 8 and in Figure 4. Let us start with columns (2) and (3), which consider R&D and non-R&D driven technological change. There is a characteristic difference between the two countries: while R&D conducting firms generate 89% of the total innovative contribution in Norway, this number is only 46% in Hungary.[64] This difference primarily results from the fact that the skill bias of non R&D based innovation is very small in Norway compared to R&D-based innovation, while the difference between the skill bias of the two types of innovation is much smaller in Hungary. In addition, R&D firms have a higher market share in Norway.

Columns (4) and (5) compare new-to-market and low-novelty innovation. In Norway, 75% of the aggregate contribution comes from new-to-market innovation, while in Hungary only 28%. The difference is mainly explained by the small prevalence of new-to-market innovation compared to Norway.

Finally, columns (6)-(8) analyze firms conducting only technical innovation, only organizational change or both type. While conducting both type of innovation is more skill biased in both countries

[63]The other two terms are very small because prior to innovation, innovative firms had a similar share in the skilled and unskilled markets, and the difference between the growth of the number and wage of low-skilled workers was very similar in innovative and non-innovative firms.

[64]These numbers are based on comparing the contributions by the two groups of firms in columns (2) and (3) of Table 8.

than conducting only one, in contrast to Norway, only technical is also highly skill-biased in Hungary. This, together with the relatively low prevalence of "both" in Hungary, explains the relatively large role of "only technical" innovation in Hungary.

These findings underline the higher importance of technology transfers—either captured by non-R&D or low-novelty innovation—in Hungary compared to Norway, where the aggregate skill bias is mainly driven by higher novelty innovation. Furthermore, technology transfers can take place by conducting only technical innovation, while in economies closer to the technology frontier, organizational changes seem to be a key driver of skill bias.

F.2 Contribution of the R&D tax credit

When estimating the effects of the R&D tax credit, we can rely on the regressions from Table 7. In particular, we can use the coefficients for the change in log H/L (0.104), log employment (0.054) and the wage premium (0.31) from Table 7.[65] The post-treatment (2006) share of treated firms after innovation was 34.6% and 38% in the college and non-college labor market, respectively. Re-running the specification in column (6) without firm fixed effects reveals that treated firms payed 4.2 percent and 1.8 percent lower wages compared to the control group.[66]

Assuming $\sigma = 2.94$ as before, the implied value of θ is 0.05, the reallocation effect is -0.14 pp, and the wage premium effect is 1.54. This yields a long-term total contribution of 1.39 percentage points, which shows that such policies can generate a large amount of skill-biased technological change that has substantial effects on the skill premium.

[65] We use the college premium effect estimated with worker fixed effects to generate conservative estimates. Using the value from column (5) yields a somewhat larger total contribution.

[66] Recall that treated firms spend less on R&D compared to control group firms, and they also pay lower wages.

Table F.1: Calculation of the Contribution of Firm-level Technological Change to Economy-wide Wage Premium

Object	Calculation
$\overline{\Delta h_j}^{inn}$	Log change in number of workers (Coefficient of "Innovation" in Table 5 col. (3)) plus log change in H/L, calculated from the change in H/L (Coefficient of "Innovation" in Table 5 col. (2)) divided by the non-innovative H/L from Table 1
$\overline{\Delta l_j}^{inn}$	Log change in number of workers (Coefficient of "Innovation" in Table 5 col. (3)) minus log change in H/L, calculated from the change in H/L (Coefficient of "Innovation" in Table 5 col. (2)) divided by the non-innovative H/L from Table 1
$\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non}$	Coefficient of "Innovation" in Table 2 column (4) but without firm fixed effects
$\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non}$	Coefficient of "Innovation" + Coefficient of "College x Innovation" in Table 2 column (4) but without firm fixed effects
ϑ_{Hjt}^{inn}	The number of college workers employed by firms with an innovation dummy=1 divided by the number of college workers employed by firms in the CIS in the 2012 wave of the CIS, weighted by CIS weights (in Norway)
ϑ_{Ljt}^{inn}	The number of non-college workers employed by firms with an innovation dummy=1 divided by the number of non-college workers employed by firms in the CIS in the 2012 wave of the CIS, weighted by CIS weights (in Norway)
$\overline{\Delta l_j}^{non}$	This can be expressed as $\ln\left(1-\vartheta_{Ljt}^{inn}\times\overline{\Delta l_j}^{inn}\right)-\ln\left(1-\vartheta_{Ljt}^{inn}\right)$, where ϑ_{Ljt}^{inn} and $\Delta\overline{l_j}^{inn}$ are calculated as described above.
$\Delta\ln\frac{\theta}{1-\theta}$	Based on Equation (13). For the change in skill ratio we take the coefficient of "Innovation x College" from column (4) of Table 2 and for the change in the skill ratio we use the coefficient of "Innovation" from column (2) of Table 5
$\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}$	Based on the coefficient of "Innovation" from Column (4) of Table 2
σ	We assume $\sigma = 2.94$ following Autor et al. (2003). In Table F.4 we show robustness to alternative values of σ

Notes: This table explains how we calculate each of the components in Equation (F.13).

Table F.2: Details of the Calculation of Contribution of Firm-level Technological Change to Economy-wide Wage Premium

Panel A: Norway

		Any	R&D	non R&D	New	Not new	Only tech.	Only org.	Both
(1)	$\overline{\Delta h_j}^{inn}$	10.86%	12.71%	7.59%	12.02%	10.37%	6.62%	9.36%	14.49%
(2)	$\overline{\Delta l_j}^{inn}$	-0.35%	-0.90%	-1.21%	-1.19%	-0.03%	1.02%	-1.85%	-3.54%
(3)	$\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non}$	5.60%	6.70%	2.70%	5.20%	2.20%	1.20%	1.60%	5.20%
(4)	$\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non}$	3.70%	4.20%	2.00%	3.80%	1.70%	0.50%	1.80%	3.80%
(5)	ϑ_{Hjt}^{inn}	61.43%	44.01%	17.09%	40.16%	20.86%	12.03%	7.56%	40.78%
(6)	ϑ_{Ljt}^{inn}	59.27%	36.24%	23.43%	34.09%	25.39%	11.63%	11.02%	37.80%
(7)	$\overline{\Delta l_j}^{non}$	0.50%	0.51%	0.37%	0.61%	0.01%	-0.14%	0.22%	2.06%
(8)	$\Delta \ln \frac{\theta}{1-\theta}$	6.39%	8.08%	2.99%	6.22%	3.86%	2.36%	3.09%	7.19%
(9)	$\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}$	-1.00%	-1.50%	0.10%	-0.60%	-0.10%	-0.70%	0.50%	-0.60%

Panel B: Hungary

		Any	R&D	non R&D	New	Not new	Only tech.	Only org.	Both
(1)	$\overline{\Delta h_j}^{inn}$	15.66%	21.41%	11.30%	34.64%	15.92%	13.93%	8.35%	24.45%
(2)	$\overline{\Delta l_j}^{inn}$	-0.43%	-1.40%	1.58%	-1.97%	0.40%	1.87%	1.62%	0.54%
(3)	$\overline{\ln w_{Hjt+1}}^{inn} - \overline{\ln w_{Hjt+1}}^{non}$	26.60%	29.90%	20.90%	31.60%	23.00%	16.00%	18.10%	20.80%
(4)	$\overline{\ln w_{Ljt+1}}^{inn} - \overline{\ln w_{Ljt+1}}^{non}$	16.60%	19.50%	13.40%	23.30%	14.00%	12.60%	10.60%	18.90%
(5)	ϑ_{Hjt}^{inn}	64.03%	30.46%	54.32%	13.40%	66.83%	8.18%	7.79%	43.75%
(6)	ϑ_{Ljt}^{inn}	66.38%	29.10%	70.17%	13.16%	86.75%	10.39%	8.37%	46.75%
(7)	$\overline{\Delta l_j}^{non}$	0.85%	0.56%	-3.85%	0.29%	-2.64%	-0.22%	-0.15%	-0.48%
(8)	$\Delta \ln \frac{\theta}{1-\theta}$	11.01%	14.16%	8.62%	16.47%	10.86%	10.44%	4.32%	14.64%
(9)	$\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}$	-0.50%	-2.10%	-0.30%	-0.50%	-1.40%	-0.10%	1.10%	-1.60%

Notes: This table shows the numerical value of the components in Equation (F.13). Each component is calculated as explained in Table F.1.

Table F.3: Decomposition Components

Panel A: Norway

		Any	R&D	non R&D	New	Not new	Only tech.	Only org	Both
(1)	Direct effect	5.61%	5.08%	0.73%	3.57%	1.15%	0.41%	0.33%	4.19%
(2)	$-\frac{1}{\sigma}\left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right) \times \left(\overline{\Delta l_j}^{inn} - \overline{\Delta l_j}^{non}\right)$	-0.01%	-0.04%	0.03%	-0.04%	0.00%	0.00%	0.02%	-0.06%
(3)	$-\left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right) \times \left(\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}\right)$	-0.02%	-0.12%	-0.01%	-0.04%	0.00%	0.00%	-0.02%	-0.02%
(4)	Wage premium effect	5.58%	4.93%	0.76%	3.49%	1.15%	0.40%	0.34%	4.11%
(5)	Reallocation effect	0.52%	0.52%	0.04%	0.34%	0.07%	0.01%	0.01%	0.37%
(6)	Total	6.10%	5.44%	0.80%	3.83%	1.22%	0.42%	0.35%	4.48%

Panel B: Hungary

		Any	R&D	non R&D	New	Not new	Only tech.	Only org	Both
(1)	Direct effect	10.07%	6.16%	6.69%	3.15%	10.37%	1.22%	0.48%	9.15%
(2)	$-\frac{1}{\sigma}\left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right) \times \left(\overline{\Delta l_j}^{inn} - \overline{\Delta l_j}^{non}\right)$	0.01%	-0.01%	-0.29%	0.00%	-0.21%	-0.02%	0.00%	-0.01%
(3)	$-\left(\vartheta_{Ljt}^{inn} - \vartheta_{Hjt}^{inn}\right) \times \left(\overline{\Delta w_{Lj}}^{inn} - \overline{\Delta w_{Lj}}^{non}\right)$	0.01%	-0.03%	0.05%	0.00%	0.28%	0.00%	-0.01%	0.05%
(4)	Wage premium effect	10.09%	6.12%	6.44%	3.15%	10.44%	1.21%	0.47%	9.19%
(5)	Reallocation effect	3.74%	2.67%	2.04%	2.01%	3.57%	0.30%	0.19%	3.25%
(6)	Total	13.83%	8.80%	8.49%	5.16%	14.00%	1.50%	0.66%	12.44%

Notes: This table shows the numerical values of the components in Equation (F.13).

Table F.4: The Contribution of Technological Change to Economy-wide College Premium over a 10-year Period: Sensitivity to σ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total contribution ($\Delta\Theta$) of:							
	Any	R&D	non R&D	New	Not new	Only tech.	Only org.	Both
Panel A: Norway								
$\sigma = 1$	9.35%	8.38%	1.48%	6.37%	2.23%	0.74%	0.74%	7.93%
$\sigma = 1.6$	7.52%	6.80%	1.08%	4.97%	1.66%	0.56%	0.52%	6.01%
$\sigma = 2.94$	6.10%	5.44%	0.80%	3.83%	1.22%	0.42%	0.35%	4.48%
$\sigma = 5$	5.45%	5.01%	0.62%	3.38%	1.00%	0.35%	0.26%	3.84%
$\sigma = 10$	4.96%	4.58%	0.52%	3.01%	0.85%	0.30%	0.20%	3.32%
Panel B: Hungary								
$\sigma = 1$	20.04%	11.28%	10.63%	6.71%	18.07%	1.96%	0.87%	16.84%
$\sigma = 1.6$	15.69%	8.37%	8.39%	4.69%	13.69%	1.54%	0.65%	12.47%
$\sigma = 2.94$	13.83%	8.80%	8.49%	5.16%	14.00%	1.50%	0.66%	12.44%
$\sigma = 5$	10.77%	5.08%	5.85%	2.40%	8.73%	1.06%	0.40%	7.52%
$\sigma = 10$	9.61%	4.30%	5.25%	1.86%	7.56%	0.95%	0.34%	6.35%

Notes: This table shows the change in the economy-wide college premium (in percentage points) due to firm-level technological change over a 10-year period based on Equation (F.13). The different columns quantify the contribution of firms conducting different forms of innovation to the aggregate college premium. We measure different forms of technological change from the detailed questionnaire of the CIS survey on firms' innovation activities. Column (1) captures the contribution of all innovative firms. Columns (2) and (3) calculate the contribution of innovators that conduct R&D and of those that do not, respectively. Columns (4) and (5) distinguish between innovators with new-to-the-market innovations, and those whose innovations are only new to the firm. Finally, columns (6), (7) and (8) calculate the contributions of firms which conducted innovations only with technical aspects (product and process), only with organizational changes, or both, respectively.