Skill-Replacing Technological Change and the Skill Premium: Theory and Evidence*

Wenbo Zhu†

July 3, 2020

Abstract

I study the effects of product innovation and process innovation on the skill premium at the industry level. I find that industries which have proportionally more firms reporting product innovations than process innovations, also tend to exhibit a higher skill premium. To better understand this phenomenon, I develop a dynamic model in which firms conduct both types of innovation endogenously. In the model, product innovation introduces new technology, which tend to complement high-skilled labour. Process innovation simplifies existing technology, which tends to complement low-skilled labour. The model generates a bi-directional relationship between the skill premium and the development of the two types of innovation. I estimate these effects based on the model using the Community Innovation Survey and the Structure of Earnings Survey from the Eurostat database. (***JEL:*** O33, E24, J24, O52)
Keywords: skill-replacing technological change, skill premium, product innovation, process innovation, Europe

*This paper is based on my PhD thesis at Queen's University. I am grateful to my supervisors, Huw Lloyd-Ellis and Sumon Majumdar for guidance and helpful discussions throughout the development of this project. I thank Charles Beach, Brant Abbott, Robert Clark, Chris Cotton, Allen Head, Beverly Lapham, Gregor Smith, Amy Sun, and Jano Zabojnik for comments and discussions, from which this draft benefited greatly. I thank an anonymous referee from the Bank of Canada for helpful feedbacks on an earlier version of this paper. I thank all the seminar participants, especially Andrea Craig, Wing Feng, Raphaelle Gauvin-Coulombe, Alex McLeod, Adugna Olani, Jason Rhinelander, David Rose, Stephen Snudden, Shaoteng Li, and Fang Xia for helpful feedback and suggestions.

†Research Institute for Global Value Chains (RIGVC), University of International Business and Economics (UIBE), Beijing, China 100029. E-mail: wenbo.zhu@uibe.edu.cn.

1 Introduction

Firms develop many kinds of technological innovations to promote their competitiveness in the market. For example, firms conduct *product innovation* to create radically novel goods and services, in order to break through new markets and reach new customer groups. Meanwhile, firms also conduct *process innovation* for their existing products to raise profit margins, usually by reducing production costs while maintaining efficiency and quality. Product innovation tends to occurs first, and as experience is accumulated, process innovation follows. New products, with embodied new technologies, are usually demanding in their implementation. As a result, high-skilled and highly educated workers are useful and often required at the initial stage of a new product. Over time, firms engage in process innovation to increase the level of specialization, which makes the production more accessible to less-expensive lower-skilled and less educated workers.[1] To this end, product innovation and process innovations could have differential impacts on labour demand.

Figure 1 plots the intensity of product and process innovation against the skill premium at the industry level. The data used here includes the Community Innovation Survey (CIS) and the Structure of Earnings Survey (SES) from the Eurostat database. More specifically, in the CIS, I observe the number of firms which report *at least one* product innovation in each industry, and the same statistic for process innovation. Meanwhile, using data in the SES, I can derive the skill premium in each industry, which is defined as the wage ratio between workers with college degrees and workers with only high school

[1] See Galor and Tsiddon (1997) and Greenwood and Yorukoglu (1997). For some recent evidence, see Figure 9 in Acemoglu and Restrepo (2018).

education.

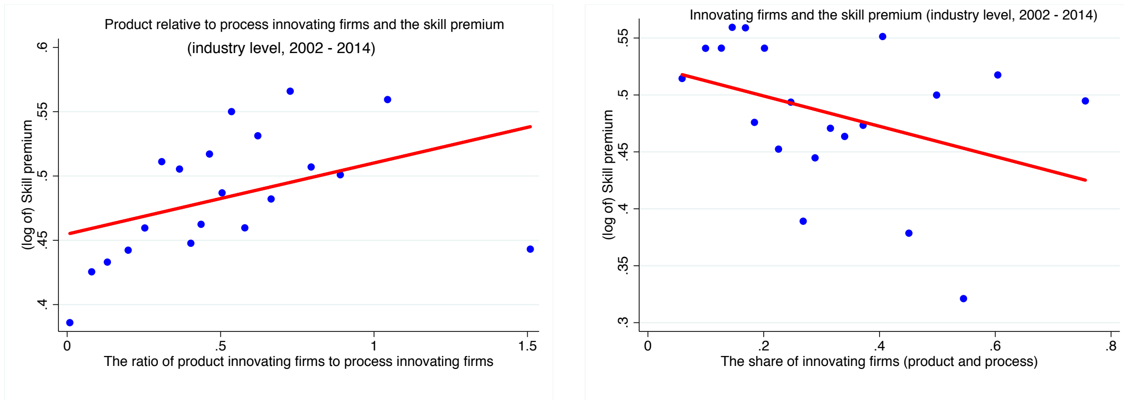
On the left panel, I plot the number of firms with at least one product innovation relative to the number of firms with at least one process innovation against the log of skill premium. For an easier interpretation, I group all my observations into 20 equal-sized bins and plot the mean value of these bins. We can see that industries with more firms reporting product innovation than process innovation also tend to have a higher skill premium. On the right panel, as a comparison, I plot the share of firms which report at least one innovation, regardless of being product or process, in an industry against the log of skill premium. Interestingly, industries with proportionally more firms reporting innovation do not tend to have a higher skill premium. From these two plots, it seems important to distinguish between the two types of innovation when discussing the impact of technological progress on the labour market.[2]

In this paper, I develop a dynamic model (a la Romer (1990)), in which firms conduct both product innovation and process innovation endogenously. Product innovation introduces new intermediate varieties into the economy. These new varieties are assumed to be non-routine, so high-skilled workers are required to implement them. After product innovation, firms can also conduct process innovation to improve upon it, by routinizing the production process and thereby reducing the skill requirements. The benefit of process innovation is that it allows firms to replace high-skilled workers with less expensive low-skilled workers.

In addition to skill-biased technological change, which is captured by the introduc-

[2]These plots are merely suggestive. The direction of causation is unclear and I am not yet controlling for other factors, including industry heterogeneity.

Figure 1: Product innovation versus process innovations



Note: on the left panel, I plot the log of skill premium against the ratio of firms reporting at least one product innovation to firms reporting at least on process innovation, at the industry level. The graph is generated using a user developed Stata®function *binscatter*. Basically, instead of raw data, I group the ratio of product and process innovating firms (i.e., the x-axis variable) into 20 equal-sized bins. I then compute the means of this ratio and the corresponding skill premium within each bin and plot them. On the right panel, I plot a similar graph, but instead of using the ratio of product to process innovations, I use the share of innovating firms in each industry regardless of the type of innovation.

tion of new products, the model also emphasizes a "deskilling" process which occurs at a later stage of the life cycle of a product/technology. I label this second type of technological change as *skill-replacing* technological change (SRTC), as it reduces the skill requirements associated with the job. Well-known examples of SRTC include assembly lines and interchangeable parts (Acemoglu (2002)). In addition, Autor (2015) discusses the idea of environmental control, and argues that engineers can sometimes simplify the environment that machines work in to enable autonomous operation. As a result, firms can disentangle different parts of a job, with machines performing the routine part and workers performing a lower skilled residual. This is another example of SRTC, since it allows firms to replace high-skilled workers with low-skilled workers, plus a suitable piece of machinery and working environment.[3]

The model demonstrates how the *composition* of technological change (i.e., product versus process) affects the skill premium by changing the relative demand for different labour skill types. In addition, the model features a bi-directional relation between the nature of innovation and the skill premium. In particular, more product innovation raises the demand for high-skilled workers and, given the labour supply, raises the skill premium. At the same time, a higher skill premium encourages firms to conduct more process innovation relative to product innovation. Thus, the equilibrium relationship between innovation composition and the skill premium reflects the net effect of these two forces.

Importantly, in addition to identify the nature of this simultaneity, the model also provides a logical underpinning for the instrumental variables needed for identification in the empirical analysis. Firms have to incur costs to engage in product and process innovations. If some of these costs affect the skill premium only through the firm's decision regarding innovations, then these cost variables can be used as instrumental variables. I conduct a comparative static analysis to illustrate this argument more formally. In the empirical analysis, I adopt industry level collaborations on innovation and economies of scale in innovation as two instruments.

The empirical analysis makes use of the CIS and the SES from the Eurostat database. The cleaned dataset covers 9 industries in 28 European countries from 2002 to 2014. The

[3] Admittedly, such improvements usually also evolve reducing the amount of human labour input. This paper chooses to focus on the skill-replacing aspect rather than the labour-saving aspect. See the first part of the literature review at the end of this section for a more developed discussion on this point.

main result I find is that the elasticity of the skill premium with respect to the intensity of product relative to process innovation ranges between 0.03 and 0.09. In comparison, the elasticity of the skill premium with respect to the quantity of labour supplied by high-skilled relative to low-skilled workers (i.e., college educated relative to high school graduates) ranges from -0.04 to -0.07. In addition, my empirical results support the bi-directional relationship featured in the model. Without correcting for the simultaneity, the estimated effect of product innovation relative to process innovation on the skill premium would be biased downwards by about 50%. Lastly, I provide estimates for several structural parameters in the model.

This paper contributes to three strands of literature. First of all, in this paper, I discuss the skill-replacing aspect of process innovations, which complements the literature on labour-saving technological changes (Frey and Osborne (2017), Acemoglu and Restrepo (2018), Hémous and Olsen (2014)). Robots and automation technologies can complement high-skilled workers and substitute for low-skilled workers. As the price of such equipment falls over time, low-skilled workers could well be pushed into lower paying manual service occupations (Autor and Dorn (2013)). Process innovation emphasizes a different and parallel channel, through which complicated jobs are standardized and can be passed on from higher skilled workers to lower skilled workers. My paper thereby provides another useful framework to think about the nature of technological changes and their impact on labour demand.

In particular, if *all* innovations are (low-skilled) labour-saving in nature, like robots and automation, then it would be difficult to reconcile the two graphs in Figure 1. In contrast, my paper focuses on the *relative* intensity of product versus process innovation, which is helpful in explaining the facts observed in the European data set.

Second, in my framework, incumbent firms, rather than entrant firms, engage in follow-up process innovations. Related literature and supporting evidence for this assumption include Foster, Haltiwanger and Krizan (2001), Bartelsman and Doms (2000), Bresnahan, Brynjolfsson and Hitt (2002), Bartelsman, Scarpetta and Schivardi (2005) and Barth et al. (2017). In this regard, my paper also relates to Acemoglu, Gancia and Zilibotti (2012), who develop a similar structure regarding product and process innovation (or product innovation and standardization in their terminology). However, they focus on the business-stealing aspect of follow-up innovations, therefore, they assume process innovation are performed by entrants. My modeling choice is motivated by the dataset used in the paper, and designed to provide a better foundation for the empirical exercise.

Lastly, using observed product and process innovation data, my model provides a novel approach to think about directly measuring skill-complementing and skill-replacing technological change separately at the industry level. In this regard, my paper contributes to the empirical literature on the Skill-Biased Technological Changes (SBTC). Related papers include Autor, Katz and Krueger (1998), Krusell et al. (2000), and etc. Combining two novel European datasets, my paper provides some direct evidence regarding both types of technological change and their impacts on the labour market. The new approach and the new evidence complement the empirical literature on SBTC.

The rest of the paper is organized as follows. Section 2 develops the model and derives the unique stationary equilibrium. Section 3 presents the empirical analysis and results.

Section 4 concludes.

2 The Model

2.1 Environment

Time is discrete. The economy is populated with high-skilled and low-skilled workers. Both types of workers share the same life-time utility function:

$$U(c_t) = \sum_{t=1}^{\infty} \beta^{t-1} \log(c_t), \tag{1}$$

where c_t denotes the final good consumed in period t. Workers choose a consumption plan to maximize life-time utility, subject to the following intertemporal budget constraint, and a no Ponzi game condition:

$$a_{t+1} = (a_t + y_t - c_t)R_{t,t+1}, \ t \in \{1, \dots, \infty\}, \tag{2}$$

where a_{t+1} and a_t denote the asset level in period $t+1$ and t, respectively. y_t denotes the workers income in period t, and $R_{t,t+1} = 1 + r_t$ denotes the gross interest rate in period t.

Labour inputs are needed in the intermediate goods sector, where wages are competitively determined. Each period, workers find the highest wage offer for their skills and supply one unit of labour endowment inelastically.

The intermediate goods sector is populated with producers using either routine or non-routine production technologies. Either high-skilled or low-skilled worker can work with routine technologies, in which case one unit of labour input generates one unit of output. In contrast, non-routine technologies require high-skilled workers and one unit of their input generates $\mu(> 1)$ units of output. Intermediate goods producers each provides one unique intermediate input for the final good sector and compete with each other in a monopolistic competitive fashion.

New intermediate goods arise as a result of *product innovation*. Innovators incur a fixed innovation cost $1/\eta$ and generate a new intermediate good. I assume with probability $\theta(> 1/2)$ the new intermediate good is non-routine and with complementary probability it is routine. The assumption allows for the possibility that product innovations are not all in the non-routine mode initially.

Non-routine intermediate producers can engage in *process innovations*, which can potentially transform their production into the routine mode. The benefit of process innovation is that it allows firms to hire low-skilled workers, which are less expensive than high-skilled workers. More specifically, at the end of each period, non-routine intermediate producers draw a producer-specific fixed cost $\tilde{\rho}_{j,t}$ from a uniform distribution with support $[\rho^0 + \lambda, \rho^1 + \lambda]$, and decide whether to conduct process innovation. The investment cost of product and process innovation are both in terms of the final good.

There is a single final good in the economy, which is produced competitively using intermediate goods, according to the following production function,

$$Y_t = N_t^{\frac{2\alpha-1}{\alpha}} \left[\sum_{j \in N_{L,t}} x_{L,j,t}^{\alpha} + \sum_{j \in N_{H,t}} x_{H,j,t}^{\alpha} \right]^{\frac{1}{\alpha}}, \alpha \in \left(\frac{1}{2}, 1 \right), \tag{3}$$

where $x_{L,j,t}$ and $x_{H,j,t}$ denote the quantities of the routine and non-routine types of intermediate inputs in period t respectively. $N_{L,t}$ and $N_{H,t}$ are measures of intermediate producers that are using the routine and the non-routine production modes respectively.[4]

The term $N_t^{\frac{2\alpha-1}{\alpha}}$ introduces a positive externality in the final good sector, when $\alpha \in (1/2, 1)$. This specification can be interpreted as "learning-by-investing": by including more intermediate goods of both types, final good producers also learn how to utilize all previous invented intermediate inputs more efficiently. It is a form of knowledge spillover, in the final good sector. This setup is adopted from Acemoglu, Gancia and Zilibotti (2012), and ensures the existence of a balanced growth path.

Lastly, the resource constraint of the economy is

$$C_t + I_t \leq Y_t, \tag{4}$$

where C_t denotes aggregate consumption, and I_t denotes the total investment in product and process innovations in period t.

The first order conditions (FOCs) for each agent of the economy are derived as fol-

[4] I use subscript L to denote routine and H to denote non-routine intermediate producers, as the minimum skill requirement for routine producers is low-skill, whereas the minimum skill requirement for non-routine producers is high-skill.

lows.[5]

- Final good producers choose intermediate inputs to minimize costs:

$$
\begin{aligned}
MP_x &= p_x, \\
x_{L,j,t} &= N_t^{\frac{2\alpha-1}{1-\alpha}} \left(\frac{1}{p_{L,j,t}}\right)^{\frac{1}{1-\alpha}} Y_t, && (5) \\
x_{H,j,t} &= N_t^{\frac{2\alpha-1}{1-\alpha}} \left(\frac{1}{p_{H,j,t}}\right)^{\frac{1}{1-\alpha}} Y_t, && (6)
\end{aligned}
$$

  where $p_{L,j,t}$ and $p_{H,j,t}$ denote the prices charged by routine and non-routine intermediate producers respectively.

- Intermediate producers choose prices to maximize profits:

$$
\begin{aligned}
p_{L,t} &= \frac{w_{L,t}}{\alpha}, && (7) \\
p_{H,t} &= \frac{w_{H,t}}{\alpha\mu}, && (8)
\end{aligned}
$$

  where $w_{L,t}$ and $w_{H,t}$ denote the market wage rates for low-skilled and high-skilled respectively.

- Workers choose a consumption stream to maximize life time utility:

$$
\frac{c_{t+1}}{c_t} = \beta(1 + r_t), \tag{9}
$$

  where r_t denotes the interest rate.

[5] The optimization problems for each agent are included in the Appendix.

2.2 Equilibrium

Given the parameters in the model, an equilibrium consists of the following objects: output, total R&D investment (on both product and process innovation), and consumption $\{Y_t, I_t, C_t\}$; measures of routine and non-routine intermediate goods producers, $\{N_{L,t}, N_{H,t}\}$; prices charged by routine and non-routine intermediate goods producers, $\{p_{L,t}, p_{H,t}\}$; wages of low-skilled and high-skilled workers, $\{w_{L,t}, w_{H,t}\}$; the interest rate $\{r_t\}$, and the probability that a non-routine intermediate good producer undertakes process innovation, $\{\gamma_t\}$, such that: (1) workers allocate themselves to the labour market which offers the highest wage for their skills; (2) workers allocate their consumption and savings to maximize their utilities; (3) intermediate good producers set prices and hire workers to maximize profits; (4) non-routine intermediate goods producers choose whether to engage in process innovation optimally; (5) final good producers choose intermediates to minimize costs and earn zero profits; (6) the research sector (product innovators) breaks even; and (7) the final good market, the intermediate goods market, the asset market, and the labour market are all clear.

The model can generate two types of labour market equilibrium, depending on parameters. In one type, all high-skilled workers are employed by non-routine intermediate producers and earn a skill premium. In contrast, in the other type of equilibrium, some high-skilled workers are employed by routine intermediate producers and earn the same wage as low-skilled workers. In the rest of the paper, I focus on the parameter space which generates the first type of equilibrium.

In this "complete sorting" type of equilibrium, the labour market clearing conditions are

$$
\begin{aligned}
L &= x_{L,t} N_{L,t}, \\
\mu H &= x_{H,t} N_{H,t},
\end{aligned}
$$

which can be re-arranged and obtain,

$$
x_{L,t} = \frac{L}{N_{L,t}}, \tag{10}
$$

$$
x_{H,t} = \frac{\mu H}{N_{H,t}}. \tag{11}
$$

Equations 10 and 11 imply that all routine (non-routine) intermediate producers would produce the same amount of output as they are symmetrical. As a result the final good production function can be simplified as

$$
Y_t = \left[{\chi_{L,t}}^{1-\alpha} L^{\alpha} + (1 - \chi_{L,t})^{1-\alpha} (\mu H)^{\alpha} \right]^{1/\alpha} N_t, \tag{12}
$$

where $\chi_{L.t} \equiv N_{L,t}/N_t$ denotes the share of routine intermediate producers in period t. It is also useful to define $\hat{y}_t = Y_t/N_t$.

From labour market clearing conditions, and using Equation 12, I can solve for the wages as

$$w_{L,t} = \alpha \hat{y}_t^{1-\alpha} \left(\frac{\chi_{L,t}}{L}\right)^{1-\alpha} N_t, \text{ and} \tag{13}$$

$$w_{H,t} = \alpha \mu^{\alpha} \hat{y}_t^{1-\alpha} \left(\frac{1-\chi_{L,t}}{H}\right)^{1-\alpha} N_t. \tag{14}$$

As a result, the implied skill premium is

$$\frac{w_{H,t}}{w_{L,t}} = \mu^{\alpha} \left(\frac{1-\chi_{L,t}}{\chi_{L,t}} \bigg/ \frac{H}{L}\right)^{1-\alpha}, \mu > 1, \text{ and } \alpha \in \left(\frac{1}{2}, 1\right). \tag{15}$$

Note that in the model high-skilled workers are both more efficient (i.e., $\mu > 1$) and more versatile (i.e., they can perform both routine and non-routine jobs). Even if $\mu = 1$, which means high-skilled workers are as efficient as low-skilled workers, they can still earn a premium as long as they are in a relatively short supply (i.e., $\left(\frac{1-\chi_{L,t}}{\chi_{L,t}} \Big/ \frac{H}{L}\right) > 1$). This second aspect is the focus of this paper: given the supply of skills, process innovations transform non-routine jobs to routine jobs, and thereby affect the relative demand for high-skilled workers and their skill premium.

Using the wages equation 13 and 14, profits for routine and non-routine intermediate producers can be derived as

$$\pi_{L,t} = (1-\alpha) \hat{y}_t^{1-\alpha} \left(\frac{\chi_{L,t}}{L}\right)^{-\alpha}, \text{ and} \tag{16}$$

$$\pi_{H,t} = (1-\alpha) \mu^{\alpha} \hat{y}_t^{1-\alpha} \left(\frac{1-\chi_{L,t}}{H}\right)^{-\alpha}, \tag{17}$$

which imply that

$$\frac{\pi_{H,t}}{\pi_{L,t}} = \mu^{\alpha} \left(\frac{1 - \chi_{L,t}}{\chi_{L,t}} \bigg/ \frac{H}{L} \right)^{-\alpha} . \tag{18}$$

Since high-skilled workers are more efficient at non-routine jobs, they increase the profits for a non-routine intermediate producer. However, when they are in a relatively short supply, high-skilled workers receive higher wages and this reduces the profit for non-routine producers. When the second effect dominates, profits for routine producers are higher and there is an incentive for non-routine producers to conduct process innovation. In this case, routine intermediate producers represent the technological "frontier" of this economy, as they have the more widely applicable technology and also the more cost-effective technology (in per output unit terms).

Given the profits, the value of routine and non-routine producers can be represented in the following recursive form,

$$V_{L,t} = \pi_{L,t} + \frac{1}{1+r_t} V_{L,t+1}, \tag{19}$$

and

$$V_{H,t} = \pi_{H,t} + \frac{1}{1+r_t} \left[\gamma_t [V_{L,t+1} - E(\tilde{\rho}_{j,t} | \tilde{\rho}_{j,t} \leq \rho_t)] + (1-\gamma_t) V_{H,t+1} \right], \tag{20}$$

where γ_t denotes the probability of successful process innovation, and $E(\tilde{\rho}_{j,t} | \tilde{\rho}_{j,t} \leq \rho_t)$ denotes the expected cost of process innovation, where ρ_t denotes the cut-off level of the cost. Given the values of routine and non-routine intermediate producers, γ_t is determined as follows

$$\gamma_t = Pr\left\{V_{H,t} \leq V_{L,t} - \tilde{\rho}_{j,t}\right\} = \frac{V_{L,t} - V_{H,t} - \rho^0 - \lambda}{\rho^1 - \rho^0}, \quad \rho^1, \rho^0, \text{ and } \lambda > 0. \tag{21}$$

Intuitively, after realizing the process innovation cost draw, the non-routine intermediate producer compares the expected payoff of undertaking the process innovation: $V_{L,t} - \tilde{\rho}_{j,t}$ to the expected payoff of not doing so: $V_{H,t}$. The producer chooses the option with a higher expected payoff. In equilibrium, there will be a cut-off level ρ_t, and all firms drawing $\tilde{\rho}_{j,t} < \rho_t$ would pay the cost and engage in process innovation.

The decision regarding entry, or product innovation, is determined by the free-entry condition,

$$V_t = \theta V_{H,t} + (1-\theta)V_{L,t} = \frac{1}{\eta}. \tag{22}$$

There is positive spill-over in the final good sector and thereby the economy features growth. The flow of non-routine and routine producers in the economy can be represented as

$$\Delta N_{H,t} = \theta g_t N_t - \gamma_t N_{H,t}, \text{ and} \tag{23}$$

$$\Delta N_{L,t} = (1-\theta) g_t N_t + \gamma_t N_{H,t}, \tag{24}$$

where, in Equation 23, the change of non-routine intermediate producers equals the measure of non-routine producers enters directly through product innovation, subtracts those conduct process innovation and become routine producers. Similarly, in Equation 24, the change of routine intermediate producers equals to the measure of routine producers enters directly through product innovation, pluses process innovating non-routine producers. The growth rates of each type of producers are therefore

$$\frac{\Delta N_{H,t}}{N_{H,t}} = \theta g_t \frac{N_t}{N_{H,t}} - \gamma_t \tag{25}$$

$$\frac{\Delta N_{L,t}}{N_{L,t}} = (1-\theta) g_t \frac{N_t}{N_{L,t}} + \gamma_t \frac{N_{H,t}}{N_{L,t}}, \tag{26}$$

I focus on a balanced growth path equilibrium (BGP), in which the output, the consumption, the wages for high-skilled and low-skilled workers, and in particular, both the measure of non-routine and routine intermediate producers grow at the constant rate g^*. Consequently, I can get

$$N^*_{H,t} = \frac{\theta g^*}{g^* + \gamma^*} N^*_t, \tag{27}$$

$$N^*_{L,t} = \frac{(1-\theta) g^* + \gamma^*}{g^* + \gamma^*} N^*_t, \tag{28}$$

and derive

$$\chi^*_L = \frac{(1-\theta) g^* + \gamma^*}{g^* + \gamma^*}, \tag{29}$$

where the asterisk denotes steady state variables. The steady states can be summarized by 9 equations and 9 unknowns. The steady state versions of these equations are provided in the Appendix. In particular, the steady state equilibrium can be solved for the rate of product innovation (g^*) and the rate of process innovations (γ^*), using steady state versions of Equation 22 and 21:

$$\frac{1}{\eta} = \theta V_H^* + (1-\theta)V_L^*, \tag{30}$$

$$\gamma^* = \frac{V_L^* - V_H^* - \rho^0 - \lambda}{\rho^1 - \rho^0}. \tag{31}$$

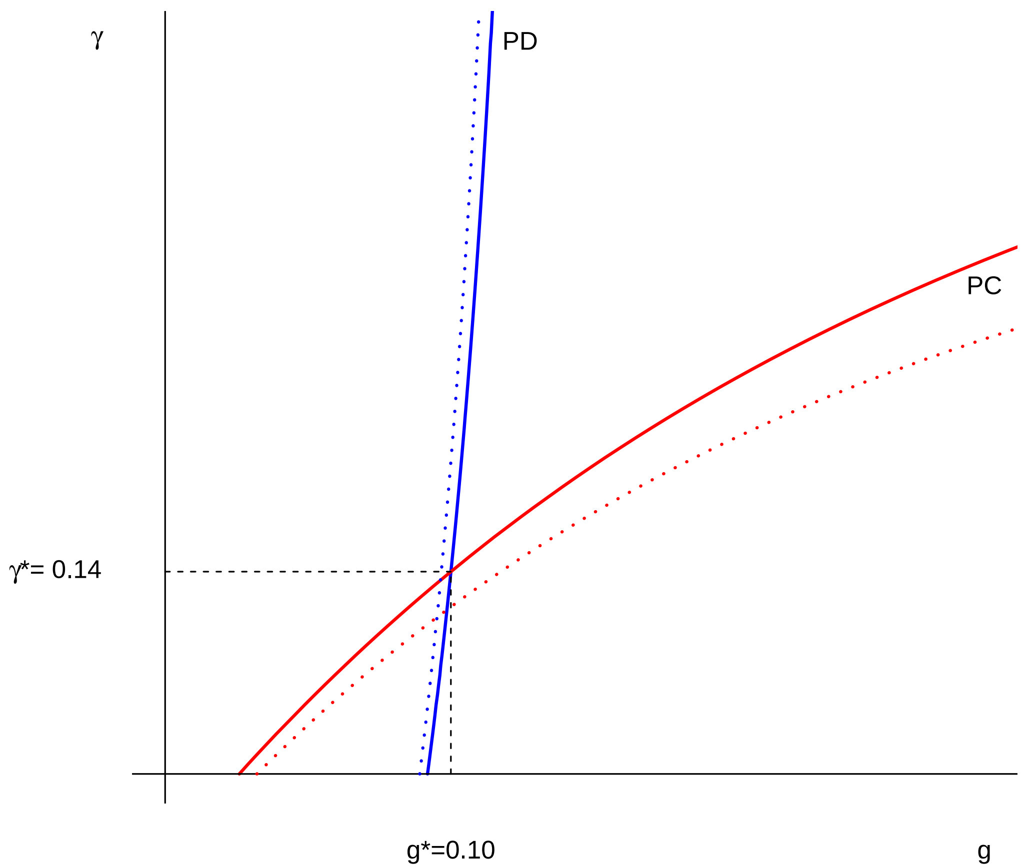
Figure 2 provides a numerical illustration of the equilibrium. The blue curve labeled as PD denotes the free entry condition (Equation 30), and the red curve labeled as PC denotes the equilibrium condition for process innovation (Equation 31).

Both curves reflect a positive relationship between g_t and γ_t. For the PD curve, when g_t increases, the labour market becomes tighter, which drives up the wages and reduces the value of product innovation. To satisfy the free entry condition, γ_t must increase. On the other hand, for the PC curve, when g_t increases, proportionally more non-routine intermediate producers enters the economy, which raises the relative demand for high-skilled workers and thereby raises the value difference between routine and non-routine producers (i.e., $V_{L,t} - V_{H,t}$ increases). As a result, non-routine intermediate producers would have stronger incentives to engage in process innovation, and thereby γ_t increases.

To explore the implications of the model and also to provide the groundwork for the subsequent empirical analysis, I conduct a comparative static analysis for an increase in the cost of conducting process innovations. More precisely, I increase the value of λ, which shifts the cost distribution of conducting process innovation to the right. Thus, in expectation it becomes more costly to engage in process innovation.

The dotted lines in Figure 2 illustrate this exercise numerically. When λ increases, the

Figure 2: Equilibrium and Comparative Static Analysis of an increase in the cost of conducting process innovations



Note: the blue curve denoted "PD" is Equation 30, whereas the red curve denoted "PC" is Equation 31. The parameters are $\beta = 0.975$, $\alpha = 0.9$, $\theta = 0.9$, $L = 5$, $H = 2$, $\mu = 1.01$, $\eta = 0.2$, $\rho^0 = 1$, $\rho^1 = 1.5$, and $\lambda = 1$. The dashed lines report the results of a comparative analysis when λ increases to 1.25. An increase in λ implies that it is more costly to conduct process innovation.

PD line shifts to the left, since for each value of γ_t, it is now less appealing to engage in product innovation. On the other hand, the PC line shifts down because for each value of γ_t, the growth rate g_t must be higher to generate a larger incentive for process innovation (i.e., to raise $V_{L,t} - V_{H,t}$), in the presence of a higher process innovation cost. Consequently, the increase in λ reduces the proportion of firms conducting process innovation, and the intensity of product innovation as well.

This exercise implies that product innovation and process innovation tend to move in the same direction. The data seem to corroborate this prediction. In the CIS, I observe that industries with a growing number of firms reporting product innovation, also tend to have a growing number of firms reporting process innovations, and vice versa. In other words, the observed changes in product innovations and process innovations by firms at the industry level are positively correlated.[6]

On the other hand, this exercise also indicates that an increase in the cost of conducting process innovation affects the skill premium, only though the decisions regarding product and process innovations. This result informs the identification strategy that I use below to address reverse causation in my estimation of the impact of the ratio of product to process innovations on the skill premium.

[6] In my data set, the correlation coefficient is 0.64 and statistically significant.

3 Empirical Analysis

3.1 Specification and Identification Strategies

The theoretical analysis in Section 2 provides a framework for the empirical analysis and illustrates the identification strategy. Firstly, to derive the specification for the empirical analysis, I take natural logarithm on both sides of Equation 15, to get

$$\log\left(\frac{w_{H,t}}{w_{L,t}}\right) = \log\mu^{\alpha} + (1-\alpha)\log\left(\frac{N_{H,t}}{N_{L,t}}\right) - (1-\alpha)\log\left(\frac{H}{L}\right). \tag{32}$$

Recall that μ denotes the productivity efficiency of high-skilled workers relative to low-skilled workers, and $N_{H,t}$ and $N_{L,t}$ are the measures of non-routine and routine intermediate producers, respectively.

Second, because of the model structure, $N_{H,t}$ and $N_{L,t}$ are proportional to the aggregate skill-complementing technology and the aggregate skill-replacing technology, which can be useful. To see this interpretation, one can simply express the equilibrium production function for the final good as

$$Y_t = N_t^{\frac{2\alpha-1}{\alpha}}\left[{N_{L,t}}^{1-\alpha}L^{\alpha} + {N_{H,t}}^{1-\alpha}(\mu H)^{\alpha}\right]^{1/\alpha}. \tag{33}$$

It is straightforward to write down a CES aggregate production function like Equation 33, and derive Equation 15 and 32 directly from it. However if one were to do so, it would be less clear how to capture two different types of production technologies (i.e., the terms appear in front of H and L in the production function) with observed variables.

This is one limitation regarding the previous empirical evidence on the SBTC. My model provides a stylized interpretation of these two types of technology, using product and process innovations.

Third, given the theoretical analysis, we can see that there is a reverse causality issue when estimating Equation 32. All else equal, a higher skill premium potentially encourages process innovation rather than product innovation, so that an OLS estimation would biases downwards the coefficient on $\log\left(N_{H,t}/N_{L,t}\right)$.

To address this issue, I try to find cost variables that affect the skill premium *only* through the decision to undertake process innovations and product innovations. The logic is discussed in the comparative static analysis above: relative costs of process innovation would affect the decision of product and process innovations but can have no direct link to the skill premium. I will discuss the two instruments in detail after I discuss the data set used in the paper.

3.2 Data

I use data from the European Union's Structure of Earnings Survey (SES) and Community Innovation Survey (CIS).

The Structure of Earnings Survey (SES) is a firm-level survey regarding the level of remuneration and the individual characteristics of employees. The objective of the SES is to provide accurate and harmonized data on earnings in European countries[7] for policy-

[7]The SES is conducted in the 28 Member States of the European Union as well as candidate countries

making and research purposes. The SES covers firms with at least 10 employees operating in all areas of the economy except public administration defined in the Statistical classification of economic activities in the European Community (NACE).[8] The SES is conducted every four years starting in 2002 and there are 4 waves of SES available so far: 2002, 2006, 2010, and 2014.[9] Harmonized (i.e., industry-level aggregation) data are publicly available through the Eurostat website and are collected for this paper.

The individual characteristics collected in the survey include age, gender, occupation, highest educational level achieved, and the length of service. In this paper, I focus on three variables from the SES, the number of employees, hours worked, and the annual gross earnings (i.e., before taxes and transfers). The number of employees and hours worked capture information about the labour supply, and the annual gross earnings captures information about wages. [10]

The Community Innovation Survey (CIS) is a firm-level survey regarding different types of innovative activities. Like the SES, the CIS also covers all EU Member States, and some EU candidate countries. As in the SES, all firms with 10 or more employees in any of the Core NACE categories are included in the statistical population.[11] The CIS

---

and countries of the European Free Trade Association (EFTA). In contrast to the European Union Labour Force Survey (LFS), the SES includes the earnings information.

[8] Information on public administration as well as enterprises with less than 10 employees is also available from some countries on a voluntary basis. However, these are not the focus of this research.

[9] The response rate varies across different countries and years. While the Eurostat quality report for SES 2014 is not available at the moment of writing, the average response rate is around 80% for SES2002, SES2006, and SES2010.

[10] I choose annual gross earnings over, for example, the hourly and monthly earnings, because annual earnings data "also includes allowances and bonuses which are not paid in each pay period, such as 13th month payments or holiday bonuses". These allowances and bonuses are an important part of some high paying occupations, which are not reflected by the hourly or monthly earnings.

[11] The full list of industries and services that are included in the Core NACE categories is provided in the

is carried out roughly every two years. Starting with CIS 3, which was conducted in 2000/2001, a standard core questionnaire was developed and applied, in order to ensure comparability across countries. There are seven waves of the survey between 2000 and 2014.[12] Compiling CIS data is voluntary for the countries, which means that in different surveys years different countries (and industries) are included.[13]

To construct the data set for this paper, I choose 2002, 2006, 2010, and 2014 as my reference years. Both surveys include information at the industry level and for most European countries. I pool all four years of data and include industry, country and year dummies in this pooled data set.

Data regarding wages and labour supplies are more standard. Consistent with most papers in the literature, I measure the "low-skilled" as ISCED 1997 level 3 and 4: high school educated and the post-secondary non-tertiary educated. I define the "high-skilled" as ISCED 1997 level 5 and 6: workers with first and second stage of tertiary education (i.e., college and above).[14] Labour income data consists of annual gross incomes, including both earnings and bonuses (before taxes and transfers). The labour supply data are constructed using the number of employees and the average hours worked (including

Appendix.

[12]To avoid confusion, I use year as an indicator for each survey, as opposed to their ordinal numbers. For example, I refer CIS 3 as CIS2000. Therefore, I have CIS2000, 2004, 2006, 2008, 2010, 2012, and 2014, seven waves of surveys altogether.

[13]According to Eurostat, the general response rate of CIS "exceeds 60% with the exception of a few countries."

[14]The International Standard Classification of Education (ISCED) was designed by UNESCO in the early 1970s to serve "as an instrument suitable for assembling, compiling and presenting statistics of education both within individual countries and internationally". In the SES 2002, 2006 and 2010, the survey used ISCED 1997 Classification, and in 2014 the survey used the ISCED 2011 Classification. The main difference is that in SES 2014 the survey further distinguishes Master and Doctoral level, which does not matter for the purpose of this paper.

both regular hours and overtime) for each industry. More details regarding the data set can be found in the appendix.

For each industry, I observe the number of enterprises which reported *at least one* successful product innovation and/or process innovation during the period under review, respectively. Note that it is possible for one enterprise to report both product and process innovations. However, if an enterprise has multiple innovations belonging to the same category, the CIS only records it once.[15]

Regarding the product innovation data, in the CIS, there are actually two different criteria to qualify a product innovation: innovations that are only new to the firm, and innovations that are not only new to the firm but also new to the firm's market (though they may have already been available in other markets). I adopt the second, arguably more stringent criterion for the empirical analysis, since the more innovative a product is, the more likely it would require high-skilled workers to implement initially. If a product innovation is only new to the firm, which means it may be available to other firms nearby, and when the current firm "adopts" this product innovation, it may be a relatively mature product already, which does not require high-skilled workers in particular for its implementation.

For each reference year, I consider innovation data from the survey of the previous wave. For example, for 2002, I use the innovation data from 2000 together with the labour income and labour supply data from 2002. Similarly, for 2006 I use the innovation data from 2004, and so on. Each CIS survey covers innovations and innovative activities for

[15]Explain the alternative setup as a robustness check. The results are similar.

a three-year period before the survey reference year. For example, CIS2000 covers all the innovations from 1998 to 2000 inclusive; CIS2004 covers 2002 to 2004 inclusive, and so on. Adopting lagged labour market data is mainly due to the consideration of two factors: it likely takes time for labour demand to be affected by innovations and also for wages to be affected by changes in labour demand. The first factor could be due to the implementation lag of innovations and the second factor could be due to wage contracts and other labour market frictions.[16]

To address the issue of reverse causation, I propose two instrumental variables. The first instrument is constructed to capture how difficult to conduct product innovations *relative to* process innovation in each industry. The second instrument is constructed to capture the economy of scale associated with product innovations *relative to* process innovations.

In the CIS, firms report successful product or process innovations which could be developed by (1) "mainly your enterprise or enterprise group", (2) "your enterprises in co-operation with other enterprises or institutions", or (3) "mainly other enterprises or institutions".[17] I consider the first case as being innovation with no or little collaboration in developing the innovation, and the latter two cases as innovation with collaboration. My conjecture is that the more collaboration in product relative to process innovations, the lower the cost of engaging in product relative to process innovations.

[16] The theoretical analysis considers only the steady state, so the connection between this justification and the model is a bit loose. If I do not use the lagged labour market data, the coefficient on the innovations (i.e., β_1) becomes less significant.

[17] In this part of CIS, there is no distinguish between product innovations and "new-to-the-market" product innovations.

More specifically, I treat the share of firms that report (2) or (3) in product innovation, relative to share of firm that report (2) and (3) in process innovation as a measure of collaboration in product relative to process innovation.[18]

It is conceivable that collaboration of innovation could also to be affected by the skill premium, as human capital is an important input to R&D activities. However, the data seems to suggest otherwise: the correlation between the skill premium and the constructed measure of relative collaboration is -0.04. In contrast, the correlation between this measure of relative collaboration and the reported product to process innovation ratio is 0.56.

On the other hand, Karshenas and Stoneman (1993) and others find that many new technologies display economies of scale.[19] Accordingly, my second instrument is constructed to capture the economies of scale in product innovation relative to process innovation. In the CIS, firms report the proportion of sales that are generated by new product(s). My conjecture is that whenever the market for new products is large, it is less costly to engage in product innovation. Conversely, if the market for current products is large, then firms are more likely to conduct process innovation, in order to find more cost-effective ways to deliver current products.

In the model, non-routine producers draw a fixed cost and decide whether to engage

[18] There is one caveat regarding the data of this instrument. From 2010 to 2014, the instruction of this question changes from "select the most appropriate only" to "tick all that apply", so the answer to this question is not unique after 2010. As a result, the percentage of firms that report (2) and (3) can be larger than 1 in product or process innovation. See the summary statistics in Table 1. However, I use the *relative* collaboration of product to process innovations as the instrument, so it is less of an issue for the current exercise.

[19] See also Davies (1979), and Baptista (2000).

in process innovation or not. This fixed cost can be thought of as a reflection of the popularity of its current product. If the cost is small, which could mean that the current product is popular, the firm would be more likely to conduct process innovation. The higher the proportion of sales that is generated by product innovations, the lower the cost of product *relative to* process innovations.

The summary statistics are provided in Table 1.

3.3 Results

To begin with, I run OLS regressions, without the instruments. This set of regressions makes use of the largest possible sample and serves as a natural starting point. Table 2 reports the regression results, and each column reports the result for one specification.

The first two columns report the "unrestricted" OLS regressions. It is re-assuring to see that both product and process innovations are important for understanding the variation of the skill premium. Column (3) reports the baseline results, which supports the main implication of the model: the number of firms that report at least one product innovation relative to process innovations is positively correlated with the skill premium, controlling for the supply of skills. Column (4) reports the implication of total innovations, and the coefficient is not significant. Column (5) provides a simple robustness check to the exercise in Column (4) by restricting the sample. Overall, the results are consistent with the model predictions.

Table 3 presents the result of the instrumental variable (IV) approach, which is the main focus of the empirical analysis. As in Table 2, each column reports the result for one specification. The first column is the main IV regression, which includes both instruments as well as time, country, and industry dummies. The result suggests that the elasticity of the skill premium with respect to the reported number of product innovation relative to that of process innovation is 0.034. In contrast, the elasticity of the skill premium with respect to the quantity of labour supplied by high-skilled workers relative to low-skilled workers is -0.041 (i.e., college-educated relative to high school graduates).[20] The error terms are clustered by time and industry simultaneously, and the estimated standard errors are obtained by the bootstrap method, with 999 replications.[21]

The last column is the OLS regression, using the same sample as the regression in the first column. In comparison to the first column, the result implies that OLS estimation biases the effects of innovations downwards by approximately 50%.[22]

Column (2) and (3) report the IV result of each individual instrument. The results are in line with the results in Column (1).

Incidentally, from the IV estimation, one can infer that the implied mark-up level (i.e., $1 - 1/\alpha$) is in the range of 2.5% to 8.8%. This estimate is well within the range (and towards the low end) of a series of mark-up values estimated for the Euro area from 1981 to 2004 by Christopoulou and Vermeulen (2008). On the other hand, one can also infer

[20] The model suggests that the magnitudes of these two coefficients should be the same, which cannot be rejected at the 10% level (the p-value of the test is 0.36). Moreover, as I have two instruments, the p-value of the Sargan test (i.e., over-identification) of this regression is 0.43.

[21] All the standard errors reported in Table 3 are obtained in this fashion, unless indicated otherwise explicitly.

[22] Note that the p-value of the Durbin-Wu-Hausman test is 0.11, which rejects the OLS estimates being consistent at the 10% level.

that the implied elasticity of substitution between high-skilled and low-skilled workers (i.e., $1/(1-\alpha)$) is in the range of 11 to 40. This result is larger than most of the estimates in the literature, see Hamermesh (1993) for a survey. One possible explanation for this large estimates is that physical capital is not included in the regression.

Interestingly, by observing the coefficients on the year dummies (not reported), I infer that all else equal, the skill premium increased in Europe from 2002 to 2014, but the trend is slowing down. Lastly, given the estimates I can infer that μ is in the range between 1.4 and 1.5, which suggests that high-skilled workers are about 40 to 50 percent more productive than low-skilled workers.

One caveat for the results above is that, for each industry, I only directly observe the number of enterprises with at least one product innovations and at least one process innovations during the period under review, respectively. Since it is entirely possible for an enterprise to have more than one innovation of each type, the estimation result could be biased.[23] For this reason, I consider an alternative measure, and try to recover the potential *total* number of product and process innovations in each industry. This measure utilizes the total number of enterprises in each industry, and a conventional structure. In particular, I assume each industry operates at the steady state level and product and process innovations arrive randomly according to the steady state Poisson rates λ_{PD} and λ_{PC}, respectively. Consequently, I can recover λ_{PD} and λ_{PC}, and represent the number of product and process innovations per unit of time as $\lambda_{PD}N_i$ and $\lambda_{PC}N_i$ respectively.[24] I present the results in the Appendix, which is not materially different

[23] The direction of this bias, however, cannot be easily determined, as only the ratio of product and process innovation is considered in the regressions.

[24] More specifically, take product innovation as an example, given the Poisson rate, the probability that

from the result in Table 3.

4 Conclusion

We live in an age of fast progressing technological changes. These technological changes are designed to work with different types of workers. Therefore, understanding the relative intensity of these technological changes, like product versus process, can help us to better grasp of the changes in the composition of labour demand. Consequently, it can also help us to better understand the changes in the income distribution. In this paper, I develop a model in which both product and process innovations can be developed by firms. I use the model to generate insights regarding how the development of these two innovations are connected and also how the two innovations affect the skill premium. I also provide empirical evidence in support for these arguments.

The development of product and process innovation can be industry-specific. For example, certain industries could be well suited to develop one type of innovation but not the other. For future research, it could be helpful to zoom into specific industries and get insights on a more granular scale. On the other hand, the CIS data set includes much richer information than what has been utilized in this paper. For example, there is information on marketing innovation and organizational innovation, on financing innovative activities, and on non-innovators. It would be fruitful to explore more of this data set.

at least one product innovation arrives before some time t is $1-e^{\lambda_{PD}t}$, which equals to $N_{H,i}/N_i$. Therefore I can solve for λ_{PD} and derive the number of innovations per period of time as $\lambda_{PD}N_i$.

References

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*, 40: 7–72.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review*, 108(6): 1488–1542.

Acemoglu, Daron, Gino Gancia, and Fabrizio Zilibotti. 2012. "Competing engines of growth: Innovation and standardization." *Journal of Economic Theory*, 147(2): 570–601.

Autor, David. 2015. "Why are there still so many jobs? The history and future of workplace automation." *The Journal of Economic Perspectives*, 29(3): 3–30.

Autor, David H, Lawrence F Katz, and Alan B Krueger. 1998. "Computing inequality: have computers changed the labor market?" *The Quarterly Journal of Economics*, 113(4): 1169–1213.

Autor, H David, and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US labor market." *The American Economic Review*, 103(5): 1553–1597.

Baptista, Rui. 2000. "Do innovations diffuse faster within geographical clusters?" *International Journal of Industrial Organization*, 18(3): 515–535.

Bartelsman, Eric J, and Mark Doms. 2000. "Understanding productivity: lessons from longitudinal microdata." *Journal of Economic Literature*, 38(3): 569–594.

Bartelsman, Eric, Stefano Scarpetta, and Fabiano Schivardi. 2005. "Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries." *Industrial and Corporate Change*, 14(3): 365–391.

Barth, Erling, James C Davis, Richard B Freeman, and Andrew J Wang. 2017. "The Effects of Scientists and Engineers on Productivity and Earnings at the Establishment Where They Work." National Bureau of Economic Research.

Bresnahan, Timothy F, Erik Brynjolfsson, and Lorin M Hitt. 2002. "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence." *The Quarterly Journal of Economics*, 117(1): 339–376.

Christopoulou, Rebekka, and Philip Vermeulen. 2008. "Markups in the Euro Area and the US over the period 1981-2014." *European Central Bank Working paper series.*

Davies, Stephen. 1979. *The Diffusion of Process Innovations.* Cambridge University Press.

Foster, Lucia, John C Haltiwanger, and Cornell John Krizan. 2001. "Aggregate productivity growth: lessons from microeconomic evidence." In *New developments in productivity analysis.* 303–372. University of Chicago Press.

Frey, Carl Benedikt, and Michael A Osborne. 2017. "The future of employment: How susceptible are jobs to computerisation?" *Technological Forecasting and Social Change*, 114: 254–280.

Galor, Oded, and Daniel Tsiddon. 1997. "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, 87(3): 363–82.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. "1974." Vol. 46, 49–95, Elsevier.

Hamermesh, Daniel S. 1993. *Labour Demand.* Princeton University Press.

Hémous, David, and Morten Olsen. 2014. "The rise of the machines: Automation, horizontal innovation and income inequality."

Karshenas, Massoud, and Paul L Stoneman. 1993. "Rank, stock, order, and epidemic effects in the diffusion of new process technologies: An empirical model." *The RAND Journal of Economics*, 503–528.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante. 2000. "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica*, 68(5): 1029–1053.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5 pt 2).

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
w_L	957	23,074	18,492	1,291	91,345
w_H	963	34,547	26,171	1,951	151,981
L	1,544	23,648,504	78,835,560	0	1,581,406,450
H	1,654	12,152,614	36,795,105	0	427,611,541
N_L	1,295	589	2,171	0	27,179
N_H	1,288	590	2,308	0	30,944
N	1,346	2,754	8,024	0	92,488
collaboration_PC	1,172	0.21	0.41	0	3.11
collaboration_PD	1,175	0.27	0.40	0	2.64
rel. economy of scale	641	0.09	0.81	0	3.18

Note: "collaboration_PC" and "collaboration_PD" denote the percentage of firms which report some form of collaboration in developing process innovation and product innovation respectively. Some of the numbers are larger than one, which is due to the non-unique categorization since 2010. See footnote 18 for an explanation.

Table 2: OLS results

	Dependent variable: $\log\left(w_{H,i}/w_{L,i}\right)$				
	(1)	(2)	(3)	(4)	(5)
$\log(N_{H,i})$	0.0287^{**}	0.0263^{**}			
	(0.0127)	(0.0125)			
$\log(N_{L,i})$	-0.0344^{***}	-0.0461^{***}			
	(0.0130)	(0.0130)			
$\log(N_{H,i}/N_{L,i})$			0.0324^{***}		
			(0.0121)		
$\log(N_{H,i}+N_{L,i})$				-0.0047	-0.0069
				(0.0046)	(0.0051)
$\log(H_i)$		-0.0605^{***}			
		(0.0118)			
$\log(L_i)$		0.0753^{***}			
		(0.0130)			
$\log(H_i/L_i)$			-0.0656^{***}	-0.0656^{***}	-0.0654^{***}
			(0.0107)	(0.0104)	(0.0108)
Constant	0.539^{***}	0.322^{**}	0.460^{***}	0.456^{**}	0.474^{***}
	(0.030)	(0.158)	(0.015)	(0.028)	(0.032)
Observations	433	433	433	460	433
R^2	0.014	0.104	0.094	0.085	0.083
F Statistic	3.59^{***}	11.92^{***}	21.60^{***}	20.05^{***}	18.44^{***}

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors are reported in the brackets.

Table 3: Instrumental Variable results

	Dependent variable: $\log\left(w^*_{H,i}/w^*_{L,i}\right)$			
	(1)	(2)	(3)	OLS
$\log(N_{H,i}/N_{L,i})$	0.034^{**}	0.025^{***}	0.088^{**}	0.018
	(0.006)	(0.006)	(0.013)	(0.018)
$\log(H_i/L_i)$	-0.041^{*}	-0.038^{*}	-0.065^{***}	-0.035
	(0.011)	(0.012)	(0.005)	(0.026)
Constant	0.365^{***}	0.365^{***}	0.366^{***}	0.326^{***}
	(0.067)	(0.067)	(0.070)	(0.081)
Collaborativeness (IV)	Y	Y		
Economy of scale (IV)	Y		Y	
Year FE	Y	Y	Y	Y
Country FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	264	264	264	264
R^2	0.59	0.59	0.56	0.59
1^{st} stage F Stat	16.50^{***}	16.12^{***}	11.38^{***}	N/A

*p<0.1; **p<0.05; ***p<0.01.

Note: Standard errors are clustered by year and industry simultaneously. In addition, bootstrap standard errors with 999 replications are reported.

A Model Specifications and Derivations

In this section, I set up the agents' problem formally and derive key equations in the main text.

Final good producers' problem

$$\max_{x_{s,j,t}} PY_t - \sum_{j \in N_t} p_{s,j,t} x_{s,j,t},\, s \in \{H, L\} \tag{34}$$

F.O.C. $x_{s,j,t}$, I get

$$MC_{s,j,t} = MP_{s,j,t}, \tag{35}$$

$$p_{s,j,t} = N_t^{\frac{2\alpha-1}{\alpha}} \frac{1}{\alpha} \left[\sum_{j \in N_{L,t}} x_{L,j,t}^{\alpha} + \sum_{j \in N_{H,t}} x_{H,j,t}^{\alpha} \right]^{\frac{1}{\alpha}-1} \alpha x_{s,j,t}^{\alpha-1}, \tag{36}$$

which can be simplified and get Equations 5 and 6.

Intermediate producers' problem

Routine intermediate producers

$$\max_{p_{L,j,t}} p_{L,j,t} x_{L,j,t} - w_{L,t} x_{L,j,t} \tag{37}$$

F.O.C. $p_{L,j,t}$, I get

$$x_{L,j,t} + (p_{L,t} - w_{L,t}) \left[N_t^{\frac{2\alpha-1}{\alpha}} \left[\sum_{j \in N_{L,t}} x_{L,j,t}^{\alpha} + \sum_{j \in N_{H,t}} x_{H,j,t}^{\alpha} \right]^{\frac{1}{\alpha}-1} \right]^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} p^{\frac{1}{\alpha-1}-1} = 0, \tag{38}$$

which can be simplified and get Equation 7.

Non-routine intermediate producers

$$\max_{p_{H,j,t}} p_{H,j,t} x_{H,j,t} - \frac{w_{H,t}}{\mu} x_{H,j,t} \tag{39}$$

F.O.C. $p_{H,j,t}$, I get

$$x_{H,j,t} + \left(p_{H,t} - \frac{w_{H,t}}{\mu}\right)\left[N_t^{\frac{2\alpha-1}{\alpha}}\left[\sum_{j\in N_{L,t}} x_{L,j,t}^{\alpha} + \sum_{j\in N_{H,t}} x_{H,j,t}^{\alpha}\right]^{\frac{1}{\alpha}-1}\right]^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} p^{\frac{1}{\alpha-1}-1} = 0, \tag{40}$$

which can be simplified and get Equation 8.

Workers' problem

$$\max_{c_t,a_{t+1}} \sum_{t=1}^{\infty} \beta^{t-1} \log(c_t), \tag{41}$$

$$\text{s.t. } a_{t+1} = (a_t + y_t - c_t)R_{t,t+1} \tag{42}$$

Set up the Lagrangian, I get

$$\mathcal{L} = \sum_{t=1}^{\infty} \beta^{t-1} \log(c_t) + \sum_{t=1}^{\infty} \lambda_t[(a_t + y_t - c_t)R_{t,t+1} - a_{t+1}] \tag{43}$$

F.O.C. c_t and a_{t+1} I get

$$\beta^{t-1}\frac{1}{c_t} - \lambda_t R_{t,t+1} = 0, \tag{44}$$

$$\lambda_{t+1}R_{t+1,t+2} - \lambda_t = 0, \tag{45}$$

which, with $R_{t,t+1} \equiv 1 + r_t$, can be used to get Equation 9.

B Steady State Equations

The steady state for the model can be represented by the following 9 equations:

$$\chi_L^* = \frac{(1-\theta)g^* + \gamma^*}{g^* + \gamma^*}, \tag{29 revisited}$$

$$\hat{y}^* = \left[\chi_L^{*\,1-\alpha} L^\alpha + (1-\chi_{L,t}^*)^{1-\alpha}(\mu H)^\alpha\right]^{1/\alpha}, \tag{46}$$

$$\pi_L^* = (1-\alpha)\hat{y}^{*^{1-\alpha}} \left(\frac{\chi_L^*}{L}\right)^{-\alpha}, \tag{47}$$

$$\pi_H^* = (1-\alpha)\mu^\alpha \hat{y}^{*^{1-\alpha}} \left(\frac{1-\chi_L^*}{H}\right)^{-\alpha}, \tag{48}$$

$$r^* = \frac{1+g^*}{\beta} - 1, \tag{49}$$

$$V_L^* = \frac{1+r^*}{r^*}\pi_L^*, \tag{50}$$

$$V_H^* = \frac{(1+r^*)\pi_H^* + \frac{\gamma^*(V_L^* - \rho^0 - \lambda}{2}}{r^* + \frac{\gamma^*}{2}}, \tag{51}$$

$$\frac{1}{\eta} = \theta V_H^* + (1-\theta)V_L^*, \tag{30 revisited}$$

$$\gamma^* = \frac{V_L^* - V_H^* - \rho^0 - \lambda}{\rho^1 - \rho^0}. \tag{31 revisited}$$

Substituting Equations 29 to 51 into Equation 30 and Equation 31 respectively, will get the PD curve (30) and the PC curve (31) in Figure 2.

C Robustness check with recovered number of product and process innovations

Table 4: Instrumental Variable results with recovered number of product and process innovations

	Dependent variable: $\log\left(w^*_{H,i}/w^*_{L,i}\right)$			
	(1)	(2)	(3)	OLS
$\log(\chi_{PD,i}/\chi_{PC,i})$	0.032^{**}	0.024^{**}	0.083^{**}	0.018
	(0.006)	(0.006)	(0.014)	(0.018)
$\log(H_i/L_i)$	-0.042^{**}	-0.038^{*}	-0.066^{***}	-0.035
	(0.010)	(0.012)	(0.002)	(0.026)
Constant	0.367^{***}	0.366^{***}	0.370^{***}	0.327^{***}
	(0.067)	(0.067)	(0.070)	(0.082)
Collaborativeness (IV)	Y	Y		
Economy of scale (IV)	Y		Y	
Year FE	Y	Y	Y	Y
Country FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	264	264	264	264
R^2	0.59	0.59	0.56	0.59
1^{st} stage F Stat	16.62^{***}	16.24^{***}	11.48^{***}	N/A

*p<0.1; **p<0.05; ***p<0.01.

Note: Standard errors are clustered by year and industry simultaneously. In addition, bootstrap standard errors with 999 replications are reported.