Skill-Replacing Process Innovation and the Labour Market: Theory and Evidence*

Wenbo Zhu†

April 13, 2022

Abstract

I study the differential impacts of product innovation and process innovation on the labour market. Using European data from 2000 to 2018, I find that industries with proportionally more firms reporting product innovation than process innovation also tend to have a lower income share of low-skilled workers. 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 intermediate goods, which tend to require high-skilled workers to implement. Process

*This paper was previously entitled "Skill-replacing technological change and the skill premium: Theory and evidence." 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 the editor Viktoria Hnatkovska and two referees for insightful comments and suggestions, from which this draft benefits greatly. I thank Charles Beach, Brant Abbott, Robert Clark, Chris Cotton, Allen Head, Beverly Lapham, Gregor Smith, Amy Sun, and Jano Zabojnik for valuable comments and discussions. 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, Kunyu Wang, and Fang Xia for helpful feedbacks and suggestions. All errors are my own.

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

innovation simplifies existing production technologies, which increases the demand for low-skilled workers. The model generates a bi-directional relationship between the skill premium and the development of the two types of innovations. I extend the model to incorporate two industries and allow low-skilled workers to switch industry freely. I calibrate the extended model to the largest two industries in UK in 2014 and 2018 respectively. I find that the product innovation has become less costly but increasingly demanding in skills, and the costs of process innovation has increased and become more diverse among different firms. (***JEL:*** O33, E24, J24, O52)

Keywords: skill-replacing technological change, skill premium, low-skilled income share, product innovation, process innovation

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 lower-skilled and less educated workers.[1] To this end, product innovation and process innovations could have differential impacts on labour demand: Product innovation can be skill-complementing, and process innovation can be skill-replacing.

To investigate this intuition empirically, I estimate the impacts of the ratio of product innovation to process innovation on the low-skilled labour income share at the industry level. The ratio of product to process innovation reflects the relative level of skill-complementing technology to the skill-replacing technology.

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

The income share of low-skilled labour is defined as the overall income received by the low-skilled divide by the overall income received by both the low-skilled and the high-skilled. The focus of this exercise is on how the composition of innovation, rather than the amount of innovation, affects relative income share for different skill groups.

One prominent issue regarding this exercise is reverse causality. If product innovation tended to require high-skilled worker, then, all else equal, a high skill premium would discourage the R&D for product innovation. Similarly, if the size of the high-skilled workforce is relatively small, then firms would also have less incentive to develop product innovation. These are the "price effect" and the "market size effect" discussed in the directed technical change literature (Acemoglu 2002a).[2]

To address this issue, I propose an instrument which leverages the different financial constraints faced by product and process innovating firms.[3] In general, product innovation tends to be more costly and risky than process innovation. Therefore, all else equal, an industry with less financial constraint would be more likely to develop product innovation as opposed to process innovation. To construct this instrument, I follow the idea of "financial dependence" proposed by Rajan and Zingales (1998) and apply it at the industry level. In particular, I calculate the proportion of gross fixed capital formation in an industry that cannot be covered by the net operating surplus generated by the industry. The larger this proportion an industry has, the more financially constrained this industry should be. Overall, I find that an industry with proportionally more product innovation than process innovation, also tend to have a lower income share of low-skilled workers.

To further understand this result in a formal setting, I develop a dynamic model (a la Romer (1990)), in which firms conduct both product and process innovation endogenously. Product innovation introduces new intermediate varieties into the economy. These new varieties are assumed to be "non-routine", so that high-skilled workers are required to implement them. After product innovation, firms can also conduct process innovation to break the complicate production process into smaller and more manageable pieces. As a result, the product becomes "routine" and low-skilled workers can start to operate it. The benefit of

[2]A similar argument also applies to process innovation and low-skilled workers.

[3]I thank an referee for pointing me towards this direction.

process innovation is reducing the labour cost of production.

In addition to skill-biased technological change, which is captured by the introduction of new products, the model also emphasizes a “de-skilling” process which occurs at a later stage of the life cycle of a product. 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 2002b). 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.[4]

In the baseline model, I assume a complete inelastic supply of skills, and focus on the impact of innovations on the skill premium. In an extension of the model, I include two different industries and assume that high-skills are industry specific and low-skills are generic. As a result, when process innovation accelerates in one industry, it would attract low-skilled workers from the other industry and thereby the income share of low-skilled workers would increase.

I calibrate this extended model to “Manufacturing” and “Wholesale and retail trade” in UK in 2014 and 2018, respectively. These are the two largest industries in the country, both in terms of employment and the number of firms. In comparing the recovered parameters from the two years, I find that (1) product innovation has become less costly to develop; (2) new product has become increasingly more “non-routine”, which means that the skill requirement of product innovation has increased; and (3) the cost of process innovation has increased on average, and perhaps more interestingly, it also has become more diverse among different firms.

Related Literature First of all, in this paper, I discuss the skill-replacing aspect of process innovations, which complements the literature on labour-saving

[4] 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.

technological changes (See Frey and Osborne (2017), Acemoglu and Restrepo (2018), Hémous and Olsen (forthcoming) among others). 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 provides another useful perspective to think about the nature of technological changes and their impact on labour demand.

Second, my paper relates to a large literature studying the innovation behaviour of incumbents (Bartelsman and Doms (2000), Foster, Haltiwanger, and Krizan (2001), Bresnahan, Brynjolfsson, and Hitt (2002), Bartelsman, Scarpetta, and Schivardi (2005), Barth et al. (2017)). In particular, my paper 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.

Lastly, my paper contributes to the large literature on estimating the impacts of innovation and technological changes on the labour markets. See Katz and Murphy (1992), Autor, Katz, and Krueger (1998), Krusell et al. (2000) among many others. In particular, I largely follow the identification strategy of Caroli and Van Reenen (2001), in which they investigate the labour market impact of organizational innovation. On top of the lagged labour market variables, which is used in Caroli and Van Reenen (2001), I also provide an instrumental variable to help address the reverse causality issue.

The rest of the paper is organized as follows. Section 2 conducts the empirical investigation regarding the differential impacts of product and process innovation on the income share of low-skilled workers. Section 3 develops the baseline model and conducts comparative static analysis. Section 4 provides two extension of the baseline model. Section 5 conducts the calibration exercise. Section 6 concludes.

2 Skill-complementing product innovation and skill-replacing process innovation

2.1 Specification

In this section, I document the differential labour demand impacts of product innovation and process innovation. In particular, I estimate the relationship between the income share of low-skilled workers and the ratio of product innovations to process innovations at the industry level. The specification of the baseline regression is the following,

$$\frac{w_{L,i,c,t}L_{i,c,t}}{w_{L,i,c,t}L_{i,c,t} + w_{H,i,c,t}H_{i,c,t}} = \beta_0 + \beta_1 \frac{PD_{i,c,t-1}}{PC_{i,c,t-1}} + YEAR + COUNTRY + IND + \epsilon_{i,c,t}, \tag{1}$$

where $w_{L,i,c,t}$ and $L_{i,c,t}$ denote the wage and employment level of low-skilled workers in industry i, country c, and year t. $w_{H,i,c,t}$ and $H_{i,c,t}$ are defined in the same fashion for high-skilled workers. $PD_{i,c,t-1}$ and $PC_{i,c,t-1}$ denote the number of firms which reports at least one product innovation and process innovations respectively, in industry i, country c, and year $t-1$. $YEAR$, $COUNTRY$, and IND denote, respectively, the year, the country, and the industry fixed effects. My working hypothesis is that product innovations tend to be skill-complementing and process innovations tend to be skill-replacing. Therefore, I expect β_1 to be negative.

To investigate the labour market impact of innovations, one usually could estimate the relationship between the *changes* of skill income share and the innovations, which represent the *changes* in technological levels (see Caroli and Van Reenen 2001 for example). Due to data limitations however, performing such a regression at the industry level reduces the number of observations drastically. As a result, I use the ratio of PD/PC as a proxy for the *relative levels* of skill-complementing technology to skill-replacing technology prevalent in the industry.

An important assumption for this interpretation is that the change of skill-complementing technology (i.e., product innovations) is proportional to the level of it. A similar relationship is also assumed for skill-replacing technology and process innovations.[5] Note that I do not need to assume the two proportions are the same. However, if one were to include PD and PC separately in the regression above, then this interpretation would not apply, and the two sides of the regression would be "unbalanced".

I conduct a few experiments to check if the identification strategy proposed would work as intended. In particular, I also observe the number of firms that reports at least one organizational innovation (ORG) in the data, which has been shown to be skill-biased in Caroli and Van Reenen (2001). I use this measure to replace PD in the numerator and see if the result is sensible. Moreover, I also, arbitrarily, use the ratio of organizational innovation to product innovation as the explanatory variable to conduct a "placebo test".

On the other hand, there could be the issue of reverse causality in the specification above. First of all, when w_L is large, which implies the skill premium is low, firms would have less incentive to conduct process innovation, relative to product innovation. On the contrary, when L is large, which implies a large market for process innovation, firms would have more incentive to conduct process innovation, relative to product innovation. Overall, the argument is related to the "price effect" versus the "market size effect" emphasized in the directed technical change literature (Acemoglu 2002a). Note that the directions of the biases are opposite to each other, with the "price effect" would bias β_1 upwards, while the "market size effect" would bias β_1 downwards.

To address the issue of reverse causality, I construct a measure of financial constraint for each industry in each country in each year as an instrumental variable, in the spirit of Rajan and Zingales (1998). More specifically, I observe the "net operating surplus and mixed income" in the data, which is defined as the gross output of an industry, less (1) the cost of intermediate goods and services, (2) compensation of employees, (3) taxes and subsidies on production, (4) imports, and (5) consumption of fixed capital. Meanwhile, I also observe the "gross fixed capital formation" for each industry in each country in each year. I then define the financial constraint of an industry as gross fixed capital formation minuses the net operating surplus and mixed income, and then divide the difference by gross fixed capital formation. The basic idea is that the larger this fraction, the more the industry relies on the external financing and thereby more financially

[5] Note that this assumption can be supported with a typical balanced growth path equilibrium. In particular, the equilibrium in the model presented later in this paper is consistent with this assumption.

constrained.[6]

The logic of this instrument is that, when compared to process innovations, product innovations tend to be more resource intensive in general. As a result, if an industry is more financially constrained, all else equal, it would then be more likely to conduct process innovation, as opposed to product innovation.

Lastly, I adopt lagged labour market data in the regression (i.e., note the $t-1$ subscripts on PD and PC). This specification is mainly due to the consideration of two factors: it likely takes time for labour demand to be affected by innovations (i.e., the implementation lag of innovations) and also for wages and employments to be affected by changes in labour demand (i.e., wage contracts and other labour market frictions).

2.2 Data

Wages and employments data comes from the Structure of Earnings Survey (SES) and the innovation data comes from the Community Innovation Survey (CIS). Both of these surveys are firm level surveys conducted by the Eurostat, which covers most European countries. Harmonized (i.e., industry-level aggregation) data is publicly available through the Eurostat website and are collected for this paper. In both surveys, all firms with 10 or more employees in any of the Core NACE categories are included in the statistical population.

The SES is conducted every four years starting in 2002 and there are 5 waves of SES available so far: 2002, 2006, 2010, 2014, and 2018. The SES collects data on the level of remuneration and the individual characteristics of employees. 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 two variables from the SES, the number of employees, which captures the information about employment, and the annual gross earnings (i.e., including both earnings and bonuses, and before taxes and transfers), which captures

[6] The main difference from the instrumental variable constructed in Rajan and Zingales (1998) is that I do not know if all the funding in the "net operating surplus and mixed income" for an industry are readily disposable. My instrumental variable can be understood as a lower bound of the financial constraint, in the sense that the industry could be more financially constrained if not all the "net operating surplus and mixed income" are disposable.

information about wages.[7]

On the other hand, the CIS is carried out roughly every two years. Starting with CIS 3, which was conducted in 2000, a standard core questionnaire was developed and applied, in order to ensure comparability across countries. There are nine waves of the survey altogether between 2000 and 2018 available. 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 CIS-2000. Given the specification and the data availability of the SES, I collect CIS-2000, 2004, 2008, 2012, 2016 five waves of surveys. Each CIS survey covers innovations and innovative activities for a three-year period before the survey reference year. For example, CIS-2000 covers all the innovative activities from 1998 to 2000 inclusive; CIS-2004 covers those from 2002 to 2004 inclusive, and so on.[8]

To combine the two data sets, I choose 2002, 2006, 2010, 2014, and 2018 as my reference years. 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 wage and employment data from 2002. Similarly, for 2006 I use the innovation data from 2004, and so on.

Data regarding wages and employment is 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. On the other hand, 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).[9]

[7]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-skilled workers, which are not reflected by the hourly or monthly earnings.

[8]Compiling CIS data is voluntary for the countries, which means that in different surveys years different countries (and industries) are included. According to Eurostat, the general response rate of CIS "exceeds 60% with the exception of a few countries."

[9]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 starting from 2014 the survey used the ISCED 2011 Classification. The main difference is that starting from SES2014 in the survey further distinguishes Master and Doctoral level workers, which does not matter for the purpose of this paper.

On the innovation front, for each industry, I observe the number of firms which reports *at least one* successful product innovation, process innovation, and/or organizational innovation (ORG) during the period under review (i.e., a three-year window), respectively. Note that it is possible for a firm to report all types of innovations. However, if a firm has multiple innovations belonging to the same category, the CIS only records it once.[10]

In addition, regarding the product innovation data, there are actually two different criteria to qualify a product innovation in the CIS: (1) innovations that are only new to the firm (PD), and (2) innovations that are not only new to the firm but also new to the firm's market (PDM). I consider both definitions in my regressions. Arguably, the second criterion is a more stringent, and thereby the product innovation under this category should be more radical and novel. Since the more innovative a product is, the more likely it would require high-skilled workers to implement initially, I expect the coefficient associated with this second definition to be larger and/or more significant.

Lastly, the data used to calculate industry level financial dependence come from the Annual National Accounts, which is also available on the Eurostat website. To take account of the investment and development lag of innovative activities, I use the previous industry levels of financial constraint. In particular, I use information from both year $t-1$ and year $t-2$ relative to the innovation year, and I label them as FD_{t-2} and FD_{t-3}, respectively.[11] The summary statistics are provided in Table 1.

2.3 Results

Table 2 reports the results for the baseline regression and a few robustness checks. Each column reports the result for one specification. Note that year, country, and industry fixed effects are included in all the regressions.

[10] In 2008, Eurostat updated its industry classification, from the Statistical Classification of Economic Activities (NACE) Rev.1 to NACE Rev.2. I follow the correspondence table provided by Eurostat and Perani and Cirillo (2015), and convert the NACE Rev.2 industries to their NACE Rev.1 counterparts in my data.

[11] Note that the innovation data is already one period prior to the labour market data, so that the financial dependence data is even more further ahead, relative to the reference years. For example, for year 2002, the innovation data comes from 2000 and the financial dependence data comes from year 1999 and 1998.

Table 1: Summary Statistics

Statistic	N	Mean	Std Dev	Min	Max
$w_{L,t}$	1518	24438.77	19411.58	1290.97	101930.00
$w_{H,t}$	1525	36474.43	27131.83	1951.00	160201.00
$log(L_t)$	1429	10.99	1.68	4.34	15.43
$log(H_t)$	1248	10.55	1.79	2.13	14.70
PD_{t-1}	1501	1103.34	3067.55	0.00	30944.00
PDM_{t-1}	1471	552.17	1610.85	0.00	20577.48
PC_{t-1}	1550	1101.78	2962.98	0.00	29247.00
ORG_{t-1}	1608	1350.91	3367.38	0.00	33044.00
N_{t-1}	1659	4751.02	10226.66	0.00	92488.00
FD_{t-2}	2720	-1.23	6.26	-189.71	23.43
FD_{t-3}	2718	-1.04	3.98	-57.58	130.84

Data Source: Eurostat
Note (1): N_{t-1} denotes the total number of firms in an industry in period $t-1$. It will be used as a control variable for robust checking purposes.
Note (2): In some country some year, a few industries exhibit rather large levels of financial constraint (i.e., above 1). After inspection, I find that in most cases, these industries report a negative level of "net operating surplus and mixed income". These observations constitute a small proportion of the data used in the IV regressions (about 2.7%). In Table 9 in Appendix A, I provide estimation results with all observation with either $FD_{t-2} > 1$ or $FD_{t-3} > 1$ removed. The results are qualitatively the same.

The first column shows the baseline results, as specified in Equation 1, which indicates that an industry with proportionally more product innovation (PD) than process innovation (PC), also exhibits a lower income share for low-skilled workers. In particular, a one unit increase in the ratio of PD/PC would reduce the income share of low-skilled workers by about 5.4%.

In the second column, I adopt the more stringent definition of product innovations, for which has to be "new to the market" (PDM). As expected, under this definition, product innovations tend to be more novel and thereby the effect on the demand for high-skilled workers are stronger. In particular, a one unit increase in the ratio of PDM/PC would reduce the income share of low-skilled workers by about 6.2%. This estimate is 14% larger than that of the previous "less novel" product innovation.

As explained in Section 2.1, an important assumption made here is that innovation flows are proportional to the level of technologies. Under this assumption PD/PC can be interpreted as the relative *level* of skill-complementing technology to the *level* of skill-replacing technology. To verify this interpretation, in the third regression, I replace product innovations with organizational innovations (ORG), which is shown to be skill-biased in Caroli and Van Reenen (2001). Reassuringly, the coefficient is also negative and significant at the 5% level.

To check for the possibility that the verification in regression (3) was simply a coincidence, I arbitrarily use $(ORG/PD)_{t-1}$ as the explanatory variable in regression (4), and we can see that the coefficient is neither statistically significant nor very different from zero.

In regression (5) and (6), I add a control for the (log of) total number of firms in each industry, and the results are similar to regression (1) and (2).

In the last two columns (7) and (8), I repeat the regressions in (1) and (2) respectively, but with three-way clustered standard errors (i.e., year-country-industry). We can see that both coefficients of interest are still significant at the 10% level.

In Table 3, I report more results for robustness checks. In particular, as there is a sizable portion of observations coming from emerging economies, I consider interacting the $YEAR$ dummy with the $COUNTRY$ dummy. The rationale is that some country specific characteristics could lead the industries to have both a high level of product innovation and a high level of high-skilled employment. These characteristics could change over time, especially for emerging economies. For example, the initially low level of intellectual property protection and level of education would both increase overtime as an emerging economy develops. The results are reported in column (1) and (2). Moreover, I also include the $YEAR \times INDUSTRY$ interaction dummies in addition to the year-country interaction dummies considered above, and the results are reported in column (3) and (4). The baseline results seem to survive with these additional dummies.

In Table 4, I report the results from the instrumental variable regressions. Recall that the instrumental variable is the financial constraint at the industry level, which is defined as the share of gross fixed capital formation that is not covered by the net operating surplus. The larger the number is, the more the industry depends on external financing. As product innovations tend to be both

Table 2: Main Results

	Dependent variable: $\frac{w_L L}{w_L L + w_H H}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(PD/PC)_{t-1}$	-0.0539***				-0.0519***		-0.0539*	
	(0.0135)				(0.0135)		(0.0233)	
$(PDM/PC)_{t-1}$		-0.0617***				-0.0588***		-0.0617*
		(0.0173)				(0.0173)		(0.0234)
$(ORG/PC)_{t-1}$			-0.0179**					
			(0.0075)					
$(ORG/PD)_{t-1}$				-0.0010				
				(0.0052)				
$\log(Total)_{t-1}$					0.0203**	0.0126		
					(0.0086)	(0.0094)		
YEAR	Y	Y	Y	Y	Y	Y	Y	Y
COUNTRY	Y	Y	Y	Y	Y	Y	Y	Y
INDUSTRY	Y	Y	Y	Y	Y	Y	Y	Y
Constant	0.5879***	0.5679***	0.5635***	0.5370***	0.4405***	0.4771***	0.5879***	0.5679***
	(0.0136)	(0.0096)	(0.0112)	(0.0093)	(0.0641)	(0.0699)	(0.0200)	(0.0088)
Observations	652	604	661	653	646	598	652	604
Adjusted R^2	0.806	0.823	0.807	0.803	0.810	0.826	0.805	0.823

Robust standard errors in parentheses, the last two columns are three-way clustered: year-country-industry.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: More Robustness Check Results - Time varying trends

	Dependent variable:			
	$\frac{w_L L}{w_L L + w_H H}$			
	(1)	(2)	(3)	(4)
$(PD/PC)_{t-1}$	-0.0677***		-0.0464***	
	(0.0149)		(0.0167)	
$(PDM/PC)_{t-1}$		-0.0841***		-0.0582***
		(0.0207)		(0.0213)
$YEAR$	Y	Y	Y	Y
$COUNTRY$	Y	Y	Y	Y
$INDUSTRY$	Y	Y	Y	Y
$Y \times COUNTRY$	Y	Y	Y	Y
$Y \times INDUSTRY$			Y	Y
Constant	0.6017***	0.5801***	0.5810***	0.5664***
	(0.0148)	(0.0112)	(0.0164)	(0.0115)
Observations	651	603	651	603
Adjusted R^2	0.814	0.834	0.822	0.844

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

more risky and costly in comparison to process innovations, the more financial dependent an industry is, the more likely the industry would develop process innovations than product innovations, all else equal.

It usually takes time to develop innovations, so the previous years financial dependence variables are used. For example, for reference year 2002, I used innovation data from year 2000, so that the financial dependence data used are from 1999 and 1998. Column (1) reports the result of IV regression with financial dependence one year prior to the innovation year (i.e., in the case of year 2000 innovations, the financial dependence of year 1999 is used). We can see that the absolute value of the coefficient becomes larger, so that direction of the bias is downwards. In Column (2), I use the financial dependence two years prior to the innovation year instead (i.e., in the case of year 2000 innovations, the financial dependence of year 1998 is used). As is shown by the Kleibergen-Paap rk statistic (Kleibergen and Paap 2006), the financial dependence two year prior to the innovations can be a weak instrument.

In Column (3), I use the two instruments jointly and conduct an overidentification test. The p-value of the test is approximately 0.46, which suggests the instruments are valid. In Column (4), I add an additional control for the (logged) total number of firms and the results are largely unaffected. In all the IV regressions, both year and country fixed effects are included and all the standard errors are two-way clustered by year and country.

In summary, I find that industries with more product innovations, relative to process innovations, also tend to have a lower income share of low-skilled workers. This finding is in support of the hypothesis noting that product innovation tends to be skill-complementing, while process innovation tends to be skill-replacing. In the next section, I develop a growth model with endogenous decision on product and process innovations and further investigate the interaction between the innovations and the labour market.

3 Baseline model

In this section, I present an endogenous growth model with decisions on product and process innovations. I first lay out the basic environment, and then I discuss the two possible equilibrium outcome on the labour market. I define the com-

Table 4: IV Regression Results: Financial dependence

	Dependent variable: $\frac{w_L L}{w_L L + w_H H}$			
	(1)	(2)	(3)	(4)
$(PD/PC)_{t-1}$	-1.5263***	-1.0405*	-1.1078***	-1.1149***
	(0.2164)	(0.5934)	(0.5445)	(0.4318)
$\log(Total)_{t-1}$				0.1112***
				(0.0232)
YEAR	Y	Y	Y	Y
COUNTRY	Y	Y	Y	Y
Instrument(s)	FD_{t-2}		FD_{t-2}	FD_{t-2}
		FD_{t-3}	FD_{t-3}	FD_{t-3}
Constant	3.1150***	2.2918**	2.4058**	1.4346**
	(0.2313)	(1.1182)	(1.0359)	(0.6374)
Observations	631	631	631	625
rk statistic	21.706	2.323	25.409	23.487
overid p-val	NA	NA	0.4665	0.9617
Adjusted R^2	-5.517	-2.182	-2.563	-1.969

Two-way clustered standard errors in parentheses (i.e., year-country)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Each column corresponds to one specification. In Column (1) I use FD_{t-2} as the instrument. In Column (2), I use FD_{t-3} as the instrument. In Column (3) and (4), I use both FD_{t-2} and FD_{t-3} as instruments.

petitive general equilibrium and solve for the balanced growth path equilibrium. Lastly, I conduct several cases of comparative static analysis.

This baseline model focuses on the short run effects by assuming a completely inelastic labour supply, in which case the effects of a changing labour demand (due to innovations) would be reflected solely in changes of the skill premium. In the next section, I explore two extensions of the baseline model. In the first extension, I explore the impact of using high-skilled workers, instead of final goods, as inputs to R&D on product and process innovation. This extension is mainly used to discuss the robustness of the baseline results. In the second extension, I extend the model to have two industries, and relax the assumption on the labour supply side by allowing low-skilled workers to move freely between the two industries. The second extension is used to discuss how changes of innovations in one industry affect the change in labour income share in both industries.

3.1 Environment

Time is discrete. The economy is populated with high-skilled and low-skilled two types of workers. The supply of each type of worker is fixed and denoted by H and L, respectively. Workers supply their labour in the intermediate goods production sector. There are two modes of production for intermediate goods: the routine mode and the non-routine mode. In the routine mode of production, either high-skilled or low-skilled labour can be used, and one unit of labour input generate one unit of output regardless of the skill level. In contrast, the non-routine mode of production requires high-skilled workers, and one unit of high-skilled labour input generates $\mu(\geq 1)$ units of output. Each intermediate goods producer provides one unique intermediate input variety and competes with each other in a monopolistic competitive fashion.

There is a single final good Y_t in the economy, which is produced competitively using intermediate goods. More specifically, the production function of the final good takes the following constant elasticity of substitution (CES) form:

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

where $x_{L,j,t}$ denotes the quantity of routine intermediate good j used in the production of the final good in period t. Note that the routine mode of production has a lower level of skill requirement and thereby the subscript L. Accordingly, $N_{L,t}$ denotes the measure of routine mode intermediate producers in period t. Similarly, $x_{H,j,t}$ and $N_{H,t}$ are defined likewise for the non-routine intermediate goods. N_t denotes the measure of all intermediate goods producers in period t, so that $N_t \equiv N_{L,t} + N_{H,t}$ for all t. In addition, α is a measure of substitutability between different intermediate goods. With the specification in Equation 2 I effectively assume that both types of intermediate goods are equality productive in producing the final goods.

The term $N_t^{\frac{2\alpha-1}{\alpha}}$ introduces a positive externality in the final good sector, whenever $\alpha \in (1/2, 1)$. This specification can be interpreted as "learning-by-investing" in the final good production sector: by including more types of intermediate goods, final good producers also learn how to utilize all previously 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).

Final good producers choose intermediate goods to minimize costs and earn zero profit, which yields the following demand functions for routine and non-routine intermediate goods:

$$x_{L,j,t} = N_t^{\frac{2\alpha-1}{1-\alpha}} \left(\frac{1}{p_{L,j,t}}\right)^{\frac{1}{1-\alpha}} Y_t, \text{ and} \tag{3}$$

$$x_{H,j,t} = N_t^{\frac{2\alpha-1}{1-\alpha}} \left(\frac{1}{p_{H,j,t}}\right)^{\frac{1}{1-\alpha}} Y_t, \tag{4}$$

where $p_{L,j,t}$ and $p_{H,j,t}$ denote the prices of routine and non-routine intermediate good j in period t, respectively.

Intermediate producers with the same production mode compete for workers, and wages are determined competitively in the labour markets. Intermediate producers choose prices to maximize profits, which yields the following pricing rule,

$$p_{L,j,t} = \frac{w_{L,t}}{\alpha}, \text{ and } p_{H,j,t} = \frac{w_{H,t}}{\alpha\mu}, \tag{5}$$

where $w_{L,t}$ and $w_{H,t}$ denote the market wages for low-skilled and high-skilled workers in period t, respectively. As a result, all the routine mode intermediate producers in the economy are symmetric, in terms of prices, quantities and profits. Similarly, all the non-routine mode intermediate producers are also symmetric.

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 θ the new intermediate good is non-routine, and with complementary probability it is routine. When $\theta > 1/2$, this assumption implies that product innovation is skill-complementing. Meanwhile, it also allows for the possibility that not all new intermediate products start in the non-routine mode initially.

Non-routine intermediate producers can engage in *process innovations*, which can transform their production mode to routine. 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 $[\underline{\rho} + \lambda, \bar{\rho} + \lambda]$, and then decide whether to conduct process innovation. 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 investment cost of product and process innovation are both in terms of the final good.

We can write down the value for the routine and the non-routine intermediate goods producers as,

$$V_{L,t} = \pi_{L,t} + \frac{1}{1 + r_t} V_{L,t+1}, \tag{6}$$

and

$$V_{H,t} = \pi_{H,t} + \frac{1}{1 + r_t} \left[\gamma_t [V_{L,t+1} - E(\tilde{\rho}_{j,t} | \tilde{\rho}_{j,t} \leq \rho_t)] + (1 - \gamma_t) V_{H,t+1} \right]. \tag{7}$$

Note that $V_{L,t}$ and $V_{H,t}$ denote the value of routine and non-routine intermediate producer in period t, respectively. $\pi_{L,t}$ and $\pi_{H,t}$ denote the per-period profit levels of routine and non-routine intermediate producer in period t, respectively. r_t denotes the risk-free interest rate. Before drawing the process innovation cost, all the non-routine intermediate producers are symmetric in terms of their values, and henceforth there is no j subscript in $V_{H,t}$.[12] The heterogenous costs of

[12]In this sense, $V_{H,t}$ denotes the *ex ante* value for non-routine intermediate producers.

process innovation only play a role in determining which non-routine producers would engage in process innovation. Afterwards, the newly converted routine producers are also symmetric.

In Equation 7, γ_t denotes the probability of drawing a favorable cost and engaging in process innovation for a non-routine intermediate producer, and $V_{L,t+1} - E(\tilde{\rho}_{j,t}|\tilde{\rho}_{j,t} \leq \rho_t)$ denotes the expected net benefit of process innovation. With complementary probability, the producer ignores the chance of process innovation (and stays as non-routine producer).

Intuitively, after realizing the process innovation cost draw, the non-routine intermediate producer compares the payoff of undertaking the process innovation: $V_{L,t+1} - \tilde{\rho}_{j,t}$ to the expected payoff of not doing so: $V_{H,t+1}$. The producer chooses the option with a higher payoff. Consequently, γ_t is determined as follows,

$$\gamma_t = Pr\left\{V_{H,t+1} \leq V_{L,t+1} - \tilde{\rho}_{j,t}\right\} = \frac{V_{L,t+1} - V_{H,t+1} - \underline{\rho} - \lambda}{\bar{\rho} - \underline{\rho}}, \tag{8}$$

where the second half of the equation emerges due to the assumption of uniform distribution.

Overall, the change in the measure of the non-routine intermediate goods equals to the inflow because of product innovation, subtracts the outflow because of process innovation. Similarly, the measure of routine intermediate goods equals to the inflow because of product innovation, pluses the inflow because of process innovation.

$$\Delta N_{H,t} = \theta g_t N_t - \gamma_t N_{H,t}, \text{ and} \tag{9}$$

$$\Delta N_{L,t} = (1-\theta) g_t N_t + \gamma_t N_{H,t}, \tag{10}$$

where g_t denotes the growth rate in period t.

On the labour supply side, both types of workers supply their one unit of labour endowment every period inelastically, and they share the same life-time utility function:

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

where c_t denotes the final good consumed in period t and β denotes the subjective discount factor. Workers choose a consumption plan to maximize utility, subject to an intertemporal budget constraint and a No-Ponzi game condition. Workers's optimization behaviour yields the following Euler condition,

$$\frac{c_{t+1}}{c_t} = \beta(1 + r_t). \tag{11}$$

Lastly, the final good market clearing condition implies that

$$C_t + I_t = Y_t, \tag{12}$$

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

3.2 "Complete-sorting" and "pooling" labour market equilibrium

Depending on values of the parameters, the model can generate two different types of labour market equilibrium. In the first type, all high-skilled workers are employed by non-routine intermediate producers and all low-skilled workers are employed by routine intermediate producers. Moreover, high-skilled workers enjoy a skill premium. I label the first type of labour market equilibrium as "complete-sorting". In the second type of equilibrium however, some high-skilled workers takes routine positions and all workers in the economy have the same wage. I label this second type as "pooling".

Suppose the economy features the "complete-sorting" equilibrium, then the labour market clearing condition implies that

$$L = x_{L,t} N_{L,t}, \text{ and} \tag{13}$$

$$\mu H = x_{H,t} N_{H,t}, \tag{14}$$

which can be re-arranged and obtain

$$x_{L,t} = \frac{L}{N_{L,t}}, \text{ and} \tag{15}$$

$$x_{H,t} = \frac{\mu H}{N_{H,t}}. \tag{16}$$

Equation 15 implies that all routine intermediate producers would produce the same amount of output as they are symmetrical. Equation 16 implies a similar outcome for the non-routine producers. As a result, the final good production function can be simplified as

$$Y_t = \left[\chi_{L,t}^{1-\alpha} L^\alpha + (1 - \chi_{L,t})^{1-\alpha} (\mu H)^\alpha\right]^{1/\alpha} N_t, \tag{17}$$

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 $y_t \equiv Y_t/N_t$.

From labour market clearing conditions, I can solve for the equilibrium wages as

$$w_{L,t} = a y_t^{1-\alpha} \left(\frac{\chi_{L,t}}{L}\right)^{1-\alpha} N_t, \text{ and} \tag{18}$$

$$w_{H,t} = a\mu^\alpha y_t^{1-\alpha} \left(\frac{1-\chi_{L,t}}{H}\right)^{1-\alpha} N_t. \tag{19}$$

Consequently, the implied skill premium is

$$\frac{w_{H,t}}{w_{L,t}} = \mu^\alpha \left(\frac{1-\chi_{L,t}}{\chi_{L,t}} \bigg/ \frac{H}{L}\right)^{1-\alpha}, \mu > 1, \text{ and } \alpha \in \left(\frac{1}{2}, 1\right). \tag{20}$$

We can see from Equation 20 that the reason for high-skilled workers to earn a skill premium is two fold: first, they are more efficient (i.e., $\mu \geq 1$), and second, they are more versatile (i.e., they can perform both routine and non-routine jobs). In other words, 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., $\frac{1-\chi_{L,t}}{\chi_{L,t}}/\frac{H}{L} > 1$). 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.

On the other hand, when the relative demand for high-skilled workers is lower than low-skilled workers, and in particular, when the effect on the relative demand margin outweighs that on the efficiency margin, then we could have a case in which $w_{H,t}/w_{L,t} < 1$. This outcome cannot be an equilibrium as high-skilled workers could then supply their labour to the routine mode intermediate producers and earn the same wage as the low-skilled. High-skilled workers will flow to the routine mode producers, until $w_{H,t}/w_{L,t} = 1$. In this case, the economy would feature a "pooling" equilibrium.

In the rest of the paper, I focus on the "complete-sorting" type of equilibrium, which is the empirically more relevant case.

3.3 General equilibrium features "complete-sorting"

Given the parameters, a competitive "complete-sorting" 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:

- Final good producers choose intermediate goods to minimize cost and earn zero profits (Equations 3 and 4).
- Intermediate goods producers set prices and hire workers to maximize profits (Equation 5).
- Non-routine intermediate producers choose whether to engage in process innovation optimally (Equation 8).
- Workers allocate themselves to the labour market which offers the highest wages for their skills.
- Workers choose a consumption plan to maximize their utilities (Equation 11).
- Product innovators breaks even (i.e., free entry)

$$\theta V_{H,t} + (1 - \theta) V_{L,t} = \frac{1}{\eta}. \tag{21}$$

- The goods markets (Equation 12), the labour markets (Equations 13 and 14), and the asset market all clear.

3.4 Balanced growth path

There exists a balanced growth path (BGP) equilibrium, 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 same constant rate g^* (I use asterisk to denote BGP variables). Note that the share of routine intermediate producers, χ_L^*, and the skill premium,w_H^*/w_L^*, do not change over time on the BGP.

The BGP equilibrium can be solved for the rate of product innovation (g^*) and the rate of process innovations (γ^*), using steady state versions of the optimal process innovation condition (Equation 8) and the free entry condition (Equation 21) :[13]

$$\gamma^* = \frac{V_L^*(g^*, \gamma^*) - V_H^*(g^*, \gamma^*) - \underline{\rho} - \lambda}{\bar{\rho} - \underline{\rho}}. \tag{22}$$

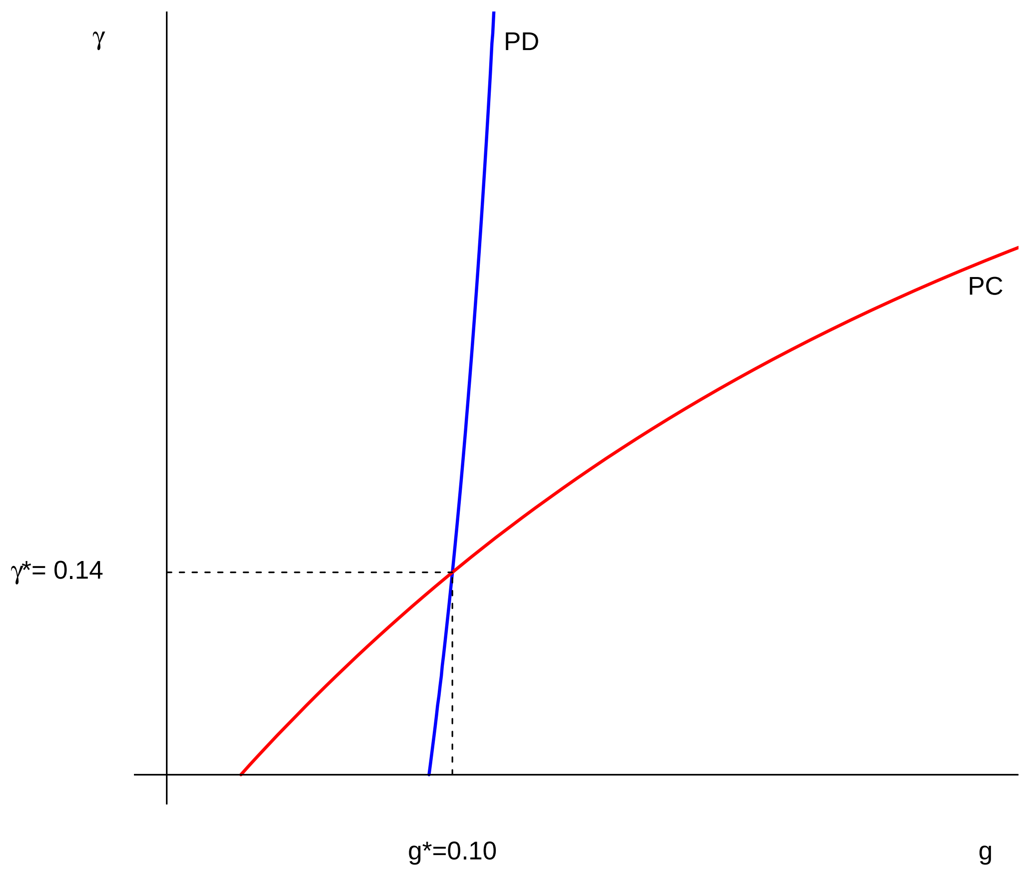
$$\frac{1}{\eta} = \theta V_H^*(g^*, \gamma^*) + (1-\theta) V_L^*(g^*, \gamma^*), \tag{23}$$

Figure 1 provides a numerical example to illustrate the equilibrium. The red curve labeled as PC denotes the equilibrium condition for process innovation (Equation 22), and the blue curve labeled as PD denotes the free entry condition (Equation 23).

Both curves reflect a positive relationship between g_t and γ_t. Regarding 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, non-routine intermediate producers must have a better chance to become routine (i.e., from the low value mode to the high value mode), so that γ_t increases. On the other hand, regarding 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.

[13]The steady state functions $V_L^*(g^*, \gamma^*)$ and $V_H^*(g^*, \gamma^*)$ and related equations and derivations are provided in Appendix B.

Figure 1: A Numerical Example for the BGP Equilibrium of the Baseline Model



Note: the blue curve denoted "PD" is Equation 23, whereas the red curve denoted "PC" is Equation 22. The parameters specified for this example 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 parameters used here are for completeness and illustrative purposes only. I calibrate an extended version of the model in Section 5.

3.5 Comparative static analysis

To explore the model's implications, I conduct three cases of comparative statics. In the first exercise, I increase λ, so that process innovations become more costly. In the second exercise, I reduce η, which implies that product innovations become more costly. In the last exercise, I increase θ, which implies that product innovations become more skill-complementing. The results of the exercises are summarized in Table 5.

Table 5: Comparative Static Analysis

	g^*	γ^*	w_H^*/w_L^*
Process innovation becomes more expensive $\lambda \uparrow$	$\downarrow$	$\downarrow$	$\uparrow$
Product innovation becomes more expensive $\eta \downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
Product innovation becomes more skill-complementing $\theta \uparrow$	$\downarrow$	$\uparrow$	$\uparrow$

In the first case, when λ increases and the cost of process innovation increases, process innovations slow down, which leads to a higher skill premium. On the other hand, the slowing down in process innovations makes the non-routine sector more congested from the firm's perspective. As a result, product innovations are discouraged and growth slows down.

In the second case, when η decreases and the cost of product innovation increases, the product innovations are discouraged and growth slows down. As a result, there is now less incentive for firms to conduct process innovation, so that γ^*decreases. The decline in demand for high-skilled workers, due to the declined product innovation, is of the first order importance, so that the skill premium decreases.

In the last case, when θ increases and the product innovations become more skill-complementing, given the supply of high-skilled workers, product innovation will slow down and g^* decreases. Consequently, there is more incentive for process innovations and thereby γ^* increases. The effect of θ dominates the effect of γ^*, which is the second order, so that the skill premium increases.

4 Extensions

4.1 High-skilled labour as input for R&D

In this extension, I consider a variation of the baseline model, in which the R&D for both product and process innovation utilize high-skilled workers, as opposed to the final goods, as the input. This extension verifies the robustness of the baseline results in an alternative environment.

First of all, I assume the aggregate R&D production function for product innovation is

$$\Delta N_t = \delta_{PD} H_{PD,t} N_t, \tag{24}$$

where ΔN_t denotes the aggregate product innovation occur in the economy in period t. δ_{PD} is a parameter denotes the productivity of high-skilled workers in developing product innovation. $H_{PD,t}$ denotes the aggregate measure of high-skilled workers engaging in product innovation. N_t denotes the total amount of intermediate varieties in period t (i.e., the amount of non-rivalry knowledge available a la Romer (1990)). Note that the marginal productivity of high-skilled workers in conducting product innovation grows at the same rate as N_t grows over time. Rearrange this production function, I get an expression for $H_{PD,t}$ as

$$H_{PD,t} = \frac{\Delta N_t}{N_t} \frac{1}{\delta_{PD}} = \frac{g_t}{\delta_{PD}}. \tag{25}$$

Accordingly, the new free entry condition becomes

$$w_{H,t} H_{PD,t} = [\theta V_{H,t} + (1-\theta) V_{L,t}] \Delta N_t, \tag{26}$$

where the left hand side denotes the total cost of product innovation in the form of total compensation to high-skilled workers engaging in product innovation in each period, whereas the right hand side denotes the total value created through product innovation in each period.

Similar to product innovation, I assume the aggregate R&D production function for process innovation is

$$\gamma_t N_{H,t} = H_{PC,t} N_t, \tag{27}$$

where $\gamma_t N_{H,t}$ denotes the aggregate measure of process innovation conducted in the economy in period t. $H_{PC,t}$ denote the aggregate measure of high-skilled workers in conducting process innovation. Note that the marginal productivity of high-skilled workers in conducting process innovation also grows at the rate g_t over time.[14] Rearrange this production function, I get an expression for $H_{PC,t}$ as

$$H_{PC,t} = \gamma_t \frac{N_{H,t}}{N_t} = \gamma_t (1 - \chi_{L,t}). \tag{28}$$

Given the labour demand for high-skilled workers in the R&D sector, the high-skilled workers in the production sector can be expressed as

$$H_{Y,t} = H - H_{PD,t} - H_{PC,t} = H - g_t/\delta_{PD} - \gamma_t (1 - \chi_{L,t}). \tag{29}$$

Accordingly, the production function of the final good becomes

$$Y_t = \left[\chi_{L,t}^{1-\alpha} L^{\alpha} + (1 - \chi_{L,t})^{1-\alpha} \mu^{\alpha} H_{Y,t}^{\alpha} \right]^{\frac{1}{\alpha}} N_t.$$

Note that the marginal productivity of high-skilled workers in the production sector also grows at the growth rate of N_t, which guarantees the proportion of high-skilled worker in each sector remains constant on the BGP.

Just as in the baseline case, I assume there is heterogeneity in the costs associated with conducting process innovation among non-routine producers. In particular, at the end of each period, each non-routine intermediate producer draws a producer-specific amount of high-skilled labour required to complete the process innovation. The labour requirements distribute uniformly between $[\underline{h} + \tau, \bar{h} + \tau]$. Upon observing the required amount of labour input, each non-routine producer decides whether to engage in process innovation or not. Let me denote the "nominal" cut-off level of labour requirement as $\tilde{h}$ and then I can derive the following cut-off condition for process innovation

$$\frac{\tilde{h}}{N_t} w_{H,t} = V_{L,t} - V_{H,t}, \tag{30}$$

where $\tilde{h}/N_t$ denotes the actual cutoff labour input requirement at time t. Note that this requirement decreases over time as high-skilled workers getting more

[14] Note that the productivity of high-skilled workers in developing process innovations is normalized to 1.

and more productive in developing process innovations.[15] Consequently, the left hand side of this equation denotes the cost of process innovation for the marginal firm in period t, whereas the right hand side of this equation denotes the benefits of process innovation.

Note that the rate of process innovation, γ_t is now defined as $\gamma_t = Prob\{V_{H,t} \leq V_{L,t} - \tilde{\tilde{h}}_t w_{H,t}/N_t\}$, where $\tilde{\tilde{h}}_t$ denotes the random draw of labour requirement from $[\underline{h} + \tau, \bar{h} + \tau]$. Substitute $w_{H,t}$ from Equation 26 and simplify, we can then obtain the following Equation 31:

$$\gamma^* = \frac{1}{\bar{h} - \underline{h}} \left(\frac{V_L^*(g^*, \gamma^*) - V_H^*(g^*, \gamma^*)}{[\theta V_H^*(g^*, \gamma^*) + (1-\theta) V_L^*(g^*, \gamma^*)]\delta_{PD}} - \underline{h} - \tau \right). \tag{31}$$

On the other hand, I can derive another equation from the fact that in equilibrium the wage of high-skilled workers in the production sector has to equal to that in the R&D sector in the steady state:

$$\alpha \mu^\alpha y^*(g^*, \gamma^*)^{*1-\alpha} \left(\frac{1 - \chi_L^*(g^*, \gamma^*)}{H_Y^*(g^*, \gamma^*)} \right)^{1-\alpha} N_t^* = [\theta V_H^*(g^*, \gamma^*) + (1-\theta) V_L^*(g^*, \gamma^*)]\delta_{PD} N_t^*, \tag{32}$$

where the LHS denotes the wage of production sector high-skilled workers, and the RHS denote the wage of R&D sector high-skilled workers.

The BGP equilibrium of the extended model can be characterized with Equations 31 and 32. I collect the derivation and function definitions in Appendix C.

I conduct a series of numerical comparative statics for the extended model. The results are presented in Table 6. First, when there is an increase in τ, which means the process innovation requires more high-skilled workers overall. In this case, process innovation is discouraged, and as a general equilibrium effect, product innovation also slows down. As the decline in process innovation dominates, the skill premium increases.

Second, when there is an increase in δ_{PD}, which means the productivity of R&D in product innovation increases and thereby the cost of product innovation

[15] I refer to $\tilde{h}$ as the "nominal" cut-off in the sense that it will always be in the rage of $[\underline{h}+\tau, \bar{h}+\tau]$ and does not change over time.

decreases, the growth rate increases. Due to a general equilibrium effect, the rate of process innovation increases as well. As the first direct impact dominates, the skill-premium increases.

Third, when there is an increase in θ, which means the product innovation become more skill-complementing, product innovation slows down, while process innovation speeds up. The effect is similar to that in the baseline case.

Lastly, when there is an increase in μ, which means the productivity of high-skilled workers in producing non-routine intermediate goods increases, the *relative* productivity of high-skilled workers in the R&D sector decreases. As a result, high-skilled workers flow out of R&D and enters the production sector. Both the product innovation and process innovation slow down, and the skill premium of high-skilled workers increases. Overall, the predictions of the extended model are qualitatively similar to the baseline case.

Table 6: Comparative Static Analysis for the Extended Model 1

	g^*	γ^*	w_H^*/w_L^*
Process innovation becomes more expensive $\tau \uparrow$	$\downarrow$	$\downarrow$	$\uparrow$
R&D productivity increases $\delta_{PD} \uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
Product innovation becomes more skill-complementing $\theta \uparrow$	$\downarrow$	$\uparrow$	$\uparrow$
Productivity in non-routine intermediates increases $\mu \uparrow$	$\downarrow$	$\downarrow$	$\uparrow$

4.2 Multiple industries

In this section, I explore another extension of the baseline model, in which there are two parallel industries for intermediate goods in the economy. To be more specific, and in line with the following calibration exercise, I label the two industries as manufacturing (M) and sales (S), respectively. Moreover, high-skilled workers are assumed to be industry-specific, and low-skilled workers are assumed to be generic and thereby can move freely between the two industries. Instead of skill-premium, this extension focuses on the change in labour income share, as a result of changes in both wages and employment.

One simplifying assumption made in this extension is that when researchers conduct product innovation, they cannot choose which industry to enter. With an exogenous probability σ_M, the new product enters manufacturing, and with complementary probability, the new product enters sales. In particular, the free entry condition becomes

$$\sigma_M \theta V_{M,H,t} + (1-\sigma_M)\theta V_{S,H,t} + (1-\theta) V_{L,t} = \frac{1}{\eta}, \tag{33}$$

where $V_{M,H,t}$ and $V_{S,H,t}$ denote the value of non-routine intermediate goods in manufacturing and sales, respectively, in period t. We can see that there is no distinguish between routine producers in the two industries in terms of values. This assumption on σ_M exogenously pins down the share of producers in two industries in the economy. In comparison to the baseline model, there is one more rate of process innovation to be determined here (i.e., one rate for manufacturing and another one for sales). The rest of the model is mechanically similar to the baseline case. The three BGP equilibrium conditions for this extension are as follows

$$\frac{1}{\eta} = \sigma_M \theta V^*_{M,H}(g^*, \gamma^*_M, \gamma^*_S) + (1-\sigma_M)\theta V^*_{S,H}(g^*, \gamma^*_M, \gamma^*_S) + (1-\theta) V^*_L(g^*, \gamma^*_M, \gamma^*_S), \tag{34}$$

$$\gamma^*_M = \frac{V^*_L(g^*, \gamma^*_M, \gamma^*_S) - V^*_{M,H}(g^*, \gamma^*_M, \gamma^*_S) - \underline{\rho}_M - \lambda_M}{\bar{\rho}_M - \underline{\rho}_M}, \tag{35}$$

$$\gamma^*_S = \frac{V^*_L(g^*, \gamma^*_M, \gamma^*_S) - V^*_{S,H}(g^*, \gamma^*_M, \gamma^*_S) - \underline{\rho}_S - \lambda_S}{\bar{\rho}_S - \underline{\rho}_S}, \tag{36}$$

where γ^*_M and γ^*_S denote the BGP equilibrium rate of process innovation for manufacturing and sales respectively. Meanwhile, I assume the cost distribution for process innovation for manufacturing uniformly distributes between $\underline{\rho}_M + \lambda_M$ and $\bar{\rho}_M + \lambda_M$, and similarly, that for sales uniformly distributes between $\underline{\rho}_S + \lambda_S$ and $\bar{\rho}_S + \lambda_S$. All the function in this extension are specified in Appendix D.

With this two-industry extension, I am interested in a scenario in which the cost of process innovation in one industry increases exogenously, and then how the skill premium and the income share of low-skilled workers in each industry would change in response to this shock. For example, what would happen if the cost of process innovation in manufacturing increases (say, due to an increase in λ_M)? It turns out that the skill premium in manufacturing would increase and the low-skilled labour income share would decrease. In the meantime, the skill premium in sales would decrease and the low-skilled income share would increase.

The intuition is that, when it is more costly to conduct process innovation in manufacturing, process innovation slows down. And it would reduce the demand for low-skilled workers in manufacturing and the wages for low-skilled workers decreases. In response to this decline in wages, low-skilled workers would flee manufacturing and enter sales. The inflow of low-skilled workers would encourage non-routine services producers to engage in process innovation to better utilize the extra supply of low-skilled workers. This result is consistent with the empirical findings in Section 2.

5 Calibration

I calibrate the extended two-industry model to the situation in UK in 2014 and 2018 respectively. In particular, I focus on "Manufacturing" (M) and "Wholesale and retail trade" (S), which are the two largest industries in UK, both in terms of employment and the number of firms.[16] There are two implicit assumptions in this exercise. First, I assume that the UK economy only has two industries and in particular, low-skilled workers move freely between manufacturing and wholesale and retail trade. Second, I assume the UK economy was in a BGP equilibrium in 2014 and in a different BGP equilibrium in 2018. The main purpose of this exercise is to use the observed labour market information and product and process innovation, to recover the unobserved costs of these innovations, and the likelihood of a new product being "non-routine".

For externally determined parameters, I choose $\alpha = 0.8$ to match the 20% markup level estimated for the Euro area between 1993 and 2004 (Christopoulou and Vermeulen 2008). I set the size of the high-skilled workforce in manufacturing to 0.87 and 0.90 for 2014 and 2018, respectively. These numbers represent millions of employees, obtained from the SES dataset. Similarly, I set the high-skilled workforce in wholesale and retail trade to 1.17 and 1.25 for 2014 and 2018, respectively. I also set the low-skilled workforce to 2.97 and 2.89 for 2014 and 2018, respectively, which reflects the sum of the low-skilled workers in manufacturing and wholesale and retail trade.[17] In addition, I set the share of the manufacturing

[16] To be more specific, "Manufacturing" refers to NACE Rev.2 Section D - Manufacturing, and "Wholesale and retail trade" refers to NACE Rev.2 Section G - Wholesale and retail trade; repair of motor vehicles and motorcycles.

[17] According to the SES, there are approximately 1.13 million low-skilled workers in manufac-

industry to 0.41 in 2014, which equals the total number of firms in manufacturing divided by the total number of firms in both industries. Similarly, I set the share of the manufacturing industry to 0.40 in 2018. Note that if the employment share is used to calculate the share of the manufacturing industry, the result is practically the same. Lastly, I set the discount rate to 0.96 and I normalize the shift parameters of process innovation costs for both industries (i.e.,λ_M,and λ_S) to zero. These parameters are summarized in the top panel of Table 7.

Table 7: Externally Determined Parameters and Calibration Results

Externally determined parameters

Parameter	2014	2018	Source
Elasticity of substitution, α	0.8	0.8	Christopoulou and Vermeulen (2008)
High-skilled workers in Manufacturing, H_M	0.87	0.90	The Structure of Earnings Survey
High-skilled workers in Wholesale and retail trade, H_S	1.17	1.25	The Structure of Earnings Survey
Low-skilled workers, L	2.97	2.89	The Structure of Earnings Survey
Share of Manufacturing firms, σ_M	0.41	0.40	The Community Innovation Survey
Process innovation cost shifting parameter for Manufacturing, λ_M	0	0	Normalization
Process innovation cost shifting parameter for Wholesale, λ_S	0	0	Normalization
The discount rate, β	0.96	0.96	Business cycle literature

Recovered parameters using the calibration procedure

Parameter	2014	2018	Target
The inverse of the cost of product innovation, η	0.183	0.218	The rate of product innovation
The probability that a new product is "non-routine", θ	0.777	0.873	The rate of PC, Skill premium and labour income share
The lower bound of process innovation cost in Manufacturing, $\underline{\rho}_M$	0.006	0.003	The rate of PC, Skill premium and labour income share
The upper bound of process innovation cost in Manufacturing, $\bar{\rho}_M$	0.018	0.610	The rate of PC, Skill premium and labour income share
The lower bound of process innovation cost in Wholesale, $\underline{\rho}_S$	0.183	0.081	The rate of PC, Skill premium and labour income share
The upper bound of process innovation cost in Wholesale, $\bar{\rho}_S$	1.311	2.401	The rate of PC, Skill premium and labour income share
The relative productivity of high-skilled workers, μ	1.363	1.238	Skill premium and labour income share

There are seven model parameters to be calibrated to match seven empirical targets. The seven empirical targets are the rate of product innovation (g), the rates of process innovation in manufacturing and wholesale and retail trade (γ_M and γ_S), and the skill premium and low-skilled labour income share in manufacturing and wholesale and retail trade ($w_{M,H}/w_{M,L}$, $w_{S,H}/w_{S,L}$, $Share_{M,L}$, and $Share_{S,L}$).[18] The first three targets come from the CIS, while the last four targets come from the SES. The values of these targets are shown in Table 8. Note that as in the model the rate of product innovation is the same for both industries, I use

turing and 1.84 million low-skilled workers in wholesale and retail trade in 2014, and in 2018, there are approximately 1.10 million low-skilled workers in manufacturing and 1.79 million low-skilled workers in wholesale and retail trade.

[18]When I construct the calibration targets, I calculate the industry specific skill premium for manufacturing and wholesale and retail trade respectively.

the average rate of product innovation in the two industries as the calibration target.

Comparing the targets's value from 2014 and 2018, we can see some interesting trends. First, the rate of product innovation increases by six percentage points, from 22% to 28%.[19] Second, the rate of process innovation also increases both in manufacturing and wholesale and retail trade. More specifically, it increases by seven percentage points from 17% to 24% in manufacturing, and by six percentage points from 8% to 14% in wholesale and retail trade. On the other hand, the skill premium in both industries decreases. In particular, it decreases about nine percent for manufacturing, and by about seven percent in wholesale and retail trade. The change in the low-skilled labour income share is much less pronounced than that in the skill premium. Nevertheless, low-skilled labour income share increases in manufacturing from 0.49 to 0.50, while it decreases in wholesale and retail trade from 0.52 to 0.51.

I present the calibration outcome next to the data moments in Table 8 for an easy comparison. Even though there are some minor disparities, the calibrated moments mostly match with the data moments, with 2018 fits better than 2014. In addition, I also include two untargeted moments in each year, which illustrate the distribution of low-skilled workers between the two industries. In particular, L_M and L_S denote the supply of low-skilled workers manufacture and in wholesale and retail trade, respectively. The calibration moments match these two moments qualitatively, with more low-skilled workers in manufacturing than the data suggests.

The recovered parameters from the calibration exercise are presented in the bottom half of Table 7. There are several interesting trends. First, the cost of product innovation reduced by about 16% between 2014 and 2018. Second, the chance that a new product requires high-skilled worker to implement (i.e., being non-routine) increased by about ten percentage points, from 77.7% to 87.3%. Third, the average cost of process innovation increased in both industries and it also becomes much more diversified. Lastly, I also find that the relative productivity of high-skilled workers decreased by about nine percent.

It seems that with product innovation becomes cheaper, but more skill de-

[19] The rate of product innovation in manufacturing is 0.28 and 0.32 for 2014 and 2018, respectively. The rate of product innovation in wholesale and retail trade is 0.16 and 0.24 for 2014 and 2018, respectively.

Table 8: Calibration Targets and Model Moments

Year	Variable	Data	Model	Targeted
2014	g	0.22	0.23	Yes
	γ_1	0.17	0.15	Yes
	γ_2	0.08	0.09	Yes
	$w_{M,H}/w_{M,L}$	1.34	1.36	Yes
	$w_{S,H}/w_{S,L}$	1.45	1.43	Yes
	$Share_{M,L}$	0.49	0.53	Yes
	$Share_{S,L}$	0.52	0.49	Yes
	L_M	1.13	1.34	No
	L_S	1.83	1.63	No
2018	g	0.28	0.28	Yes
	γ_1	0.24	0.24	Yes
	γ_2	0.14	0.14	Yes
	$w_{M,H}/w_{M,L}$	1.22	1.25	Yes
	$w_{S,H}/w_{S,L}$	1.35	1.33	Yes
	$Share_{M,L}$	0.50	0.54	Yes
	$Share_{S,L}$	0.51	0.49	Yes
	L_M	1.10	1.32	No
	L_S	1.79	1.57	No

manding, the competition for high-skilled workers is intensified and the incentive to conduct process innovation becomes larger. Despite the increasing cost, the rate of process innovation increases and thereby the skill premium decreases. The increase in process innovation is slightly more intense in manufacturing, which induces more low-skilled workers flow in from wholesale and retail trade.

6 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 the changes in the composition of labour demand. Consequently, it can also help us to better understand the changes in the income distribution and relative income share among different skill groups. In this paper, I document a new stylized fact that industries with proportionally more product innovation than process innovation also tend to have a lower income share for the low-skilled. I also develop a dynamic model to illustrate the two-way interaction between innovations and the labour markets. I calibrate an extended version of the model to match the largest two industries in the UK in 2014 and 2018, which recovers some interesting results on the distinctive changes of the costs of product and process innovation.

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. It could be fruitful to explore more of this data set.

References

Acemoglu, Daron. 2002a. "Directed Technical Change". *The Review of Economic Studies* 69 (4): 781–809.

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

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

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.

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.

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, et al. 2017. “The Effects of Scientists and Engineers on Productivity and Earnings at the Establishment Where They Work”.

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.

Caroli, Eve, and John Van Reenen. 2001. “Skill-biased organizational change? Evidence from a panel of British and French establishments”. *The Quarterly Journal of Economics* 116 (4): 1449–1492.

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.*

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, no. 3 (): 363–82.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. “1974”. In *Carnegie-Rochester conference series on public policy*, 46:49–95. Elsevier.

Hémous, David, and Morten Olsen. Forthcoming. “The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality”. *American Economic Journal: Macroeconomics.*

Katz, Lawrence F, and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors". *The Quarterly Journal of Economics* 107 (1): 35–78.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition". *Journal of Econometrics* 133 (1): 97–126.

Krusell, Per, et al. 2000. "Capital-skill complementarity and inequality: A macroeconomic analysis". *Econometrica* 68 (5): 1029–1053.

Perani, Giulio, and Valeria Cirillllo. 2015. "Matching industry classifications. A method for converting NACE Rev.2 to NACE Rev.1". *Working Paper, University of Urbino, Faculty of Economics.*

Rajan, Raghuram G, and Luigi Zingales. 1998. "Financial Dependence and Growth". *American Economic Review*: 559–586.

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

A Instrumental variable regression results without industries with extremely large levels of financial constraints

As shown in the summary statistics, in some country some year, a few industries exhibit extremely large level of financial constraints (i.e., above 1). In Table 9, I present the IV regression results with all observations with either $FD_{t-2} > 1$ or $FD_{t-3} > 1$ removed from the sample.

Table 9: IV Regression Results: Financial dependence - robustness

	Dependent variable: $\frac{w_L L}{w_L L + w_H H}$			
	(1)	(2)	(3)	(4)
$(PD/PC)_{t-1}$	-1.4523***	-1.0337*	-1.1173**	-0.9893**
	(0.2672)	(0.6237)	(0.5524)	(0.4250)
$\log(Total)_{t-1}$				0.1038***
				(0.0212)
YEAR	Y	Y	Y	Y
COUNTRY	Y	Y	Y	Y
Instrument(s)	FD_{t-2}		FD_{t-2}	FD_{t-2}
		FD_{t-3}	FD_{t-3}	FD_{t-3}
Constant	2.5667***	2.0616***	2.1625***	1.2590***
	(0.3225)	(0.7526)	(0.6666)	(0.3761)
Observations	614	614	614	608
rk statistic	24.482	1.913	14.202	16.711
overid p-val	NA	NA	0.5169	0.9070
Adjusted R^2	-4.797	-2.085	-2.548	-1.332

Two-way clustered standard errors in parentheses (i.e., year-country)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Each column corresponds to one specification. In Column (1) I use FD_{t-2} as the instrument. In Column (2), I use FD_{t-3} as the instrument. In Column (3) and (4), I use both FD_{t-2} and FD_{t-3} as instruments.

B Derivation and equations for the BGP equilibrium in the baseline model

In deriving the BGP equilibrium conditions (i.e., Equations 22 and 23), I first use Equations 9 and 10, and get

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

I then impose the BGP conditions and get

$$\theta g^* \frac{N^*}{N_H^*} - \gamma^* = g^*, \text{ and}$$
$$(1-\theta) g^* \frac{N^*}{N_H^*} + \gamma^* \frac{N_H^*}{N_L^*} = g^*,$$

which can be used to derive

$$\chi_L^* = \frac{(1-\theta)g^* + \gamma^*}{g^* + \gamma^*}. \tag{37}$$

Consequently, the final goods product function normalized by total measure of firms becomes

$$y^*(g^*, \gamma^*) = \left[\chi_L^*(g^*, \gamma^*)^{1-\alpha} L^\alpha + [1 - \chi_L^*(g^*, \gamma^*)]^{1-\alpha} (\mu H)^\alpha\right]^{\frac{1}{\alpha}}$$

in the BGP equilibrium. Then I can write down the per-period profits for routine and non-routine intermediate producers as, respectively,

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

and

$$\pi_H^*(g^*, \gamma^*) = (1-\alpha) \mu^\alpha y^*(g^*, \gamma^*)^{1-\alpha} \left(\frac{1 - \chi_L^*(g^*, \gamma^*)}{H}\right)^{-\alpha}.$$

The Euler Equation becomes

$$r^*(g^*) = \frac{1+g^*}{\beta} - 1.$$

Lastly, the value of routine intermediate producers is

$$V_L^*(g^*, \gamma^*) = \frac{1 + r^*(g^*)}{r^*(g^*)} \pi_L^*(g^*, \gamma^*),$$

and the value of non-routine intermediate producers is

$$V_H^*(g^*, \gamma^*) = \frac{[1 + r^*(g^*)]\pi_H^*(g^*, \gamma^*) + \gamma^*[V_L^*(g^*, \gamma^*) - \underline{\rho} - \lambda]/2}{r^*(g^*) + \gamma^*/2}.$$

C Derivation and equations for extension 4.1

The following derivation and equations are provided for the BGP equilibrium in the first extension (i.e., Equations 31 and 32), in which high-skilled workers, as opposed to final goods, are used as inputs for the R&D of product and process innovations.

First of all, the BGP share of routine intermediate producers is still

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

As a result, I can write down the measure of high-skilled workers in the production sector as

$$H_Y^*(g^*, \gamma^*) = H - \frac{g^*}{\delta_{PD}} - \gamma^*[1 - \chi_L^*(g^*, \gamma^*)],$$

and the final goods production function normalized by total measure of firms becomes

$$y^*(g^*, \gamma^*) = \left[\chi_L^*(g^*, \gamma^*)^{1-\alpha} L^\alpha + [1 - \chi_L^*(g^*, \gamma^*)]^{1-\alpha} \mu^\alpha H_Y^*(g^*, \gamma^*)^\alpha\right]^{\frac{1}{\alpha}}.$$

Consequently, by applying the labour market clearing condition I can derive the wage for high-skilled workers in the production sector as

$$w_{H,t}^*(g^*, \gamma^*) = \alpha\mu^\alpha y^*(g^*, \gamma^*)^{1-\alpha} \left(\frac{1 - \chi_L^*(g^*, \gamma^*)}{H_Y^*(g^*, \gamma^*)}\right)^{1-\alpha} N_t^*(g^*). \quad (38)$$

Note that by Equations 24 and 26, I can derive the wage for high-skilled workers conducting R&D for product innovation, then together with Equation 38, I can derive Equation 32.

On the other hand, the per period profits for routine and non-routine intermediate producers are, respectively,

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

and

$$\pi_H^*(g^*, \gamma^*) = (1-\alpha) \mu^\alpha y^*(g^*, \gamma^*)^{1-\alpha} \left(\frac{1 - \chi_L^*(g^*, \gamma^*)}{H_Y^*(g^*, \gamma^*)} \right)^{-\alpha}.$$

The Euler Equation is

$$r^*(g^*) = \frac{1+g^*}{\beta} - 1,$$

and lastly, the value of routine intermediate producers is

$$V_L^*(g^*, \gamma^*) = \frac{1+r^*(g^*)}{r^*(g^*)} \pi_L^*(g^*, \gamma^*),$$

and the value of non-routine intermediate producers is

$$V_H^*(g^*, \gamma^*) = \frac{2[1+r^*(g^*)]\pi_H^*(g^*, \gamma^*) + \gamma^* \left[1 + (\underline{h} + \tau)(1-\theta)\delta_{PD}\right] V_L^*(g^*, \gamma^*)}{2r^*(g^*) + \gamma^* - \gamma^*(\underline{h} + \tau)\theta\delta_{PD}}.$$

D Equations for extension 4.2

The following equations are provided for the BGP equilibrium in the second extension (i.e., Equations 34, 35, and 36), in which two separate industries, "Manufacturing" and "Sales" , are included in the baseline model.

Value of non-routine intermediate producers in Manufacturing:

$$V_{M,H}^*(g^*, \gamma_M^*, \gamma_S^*) = \frac{[1 + r^*(g^*)]\pi_{M,H}^*(g^*, \gamma_M^*, \gamma_S^*) + \gamma_M^*(V_L^*(g^*, \gamma_M^*, \gamma_S^*) - \underline{\rho}_M - \lambda_M)/2}{r^*(g^*) + \gamma_M^*/2}.$$

Value of non-routine intermediate producers in Sales:

$$V_{S,H}^*(g^*, \gamma_M^*, \gamma_S^*) = \frac{[1 + r^*(g^*)]\pi_{S,H}^*(g^*, \gamma_M^*, \gamma_S^*) + \gamma_S^*(V_L^*(g^*, \gamma_M^*, \gamma_S^*) - \underline{\rho}_S - \lambda_S)/2}{r^*(g^*) + \gamma_S^*/2}.$$

Value of routine intermediate producers (same for both industries):

$$V_L^*(g^*, \gamma_M^*, \gamma_S^*) = \frac{1 + r^*(g^*)}{r^*(g^*)}\pi_L^*(g^*, \gamma_M^*, \gamma_S^*).$$

Euler Equation:

$$r^*(g^*) = \frac{1 + g^*}{\beta} - 1.$$

Per period profits for non-routine intermediate producers in Manufacturing:

$$\pi_{M,H}^*(g^*, \gamma_M^*, \gamma_S^*) = (1 - \alpha)y^*(g^*, \gamma_M^*, \gamma_S^*)^{1-\alpha}\mu^{1-\alpha}\left(\frac{\chi_{M,H}^*(g^*, \gamma_M^*)}{H_M}\right)^{-\alpha}.$$

Per period profits for non-routine intermediate producers in Sales:

$$\pi_{S,H}^*(g^*, \gamma_M^*, \gamma_S^*) = (1 - \alpha)y^*(g^*, \gamma_M^*, \gamma_S^*)^{1-\alpha}\mu^{1-\alpha}\left(\frac{\chi_{S,H}^*(g^*, \gamma_S^*)}{H_S}\right)^{-\alpha}.$$

Per period profits for routine intermediate producers (same for both industries):

$$\pi_L^*(g^*, \gamma_M^*, \gamma_S^*) = (1 - \alpha)y^*(g^*, \gamma_M^*, \gamma_S^*)^{1-\alpha}\left(\frac{1 - \chi_{M,H}^*(g^*, \gamma_M^*) - \chi_{S,H}^*(g^*, \gamma_S^*)}{L}\right)^{-\alpha}.$$

Final goods production function normalized by the total measure of firms:

$$y^*(g^*, \gamma_M^*, \gamma_S^*) = \Big[(\mu H_M)^\alpha \chi_{M,H}^*(g^*, \gamma_M^*)^{1-\alpha} + (\mu H_S)^\alpha \chi_{S,H}^*(g^*, \gamma_S^*)^{1-\alpha}$$
$$+L^\alpha [1 - \chi_{M,H}^*(g^*, \gamma_M^*) - \chi_{S,H}^*(g^*, \gamma_S^*)]^{1-\alpha}\Big]^{1/\alpha}.$$

Share of non-routine intermediate producers in Manufacturing:

$$\chi_{M,H}^*(g^*, \gamma_M^*) = \frac{\sigma_M \theta g^*}{g^* + \gamma_M^*}.$$

Share of non-routine intermediate producers in Sales:

$$\chi_{S,H}^*(g^*, \gamma_S^*) = \frac{(1 - \sigma_M)\theta g^*}{g^* + \gamma_S^*}.$$