Canadian Labour Economics Forum

Working Paper Series

Skill-Biased Technological Change, Training, and the College Wage Premium: A Quantitative Analysis

Thomas Palmer (McMaster University)

CLEF WP #72

Skill-Biased Technological Change, Training, and the College Wage Premium: A Quantitative Evaluation*

Thomas Palmer†

May 19, 2024

Click here for most recent version

Abstract

This paper establishes that the rise in employer-provided training due to technological change has dampened the college wage premium. Using unique survey micro-data, I show that high-technology firms provide more training overall, but the gap in training participation between high- and low-skill workers is smaller within these firms. To understand the aggregate implications of these patterns, I build a quantitative model of the labor market with endogenous technology and training investments. In a counterfactual exercise, I find that the increase in the college wage premium would be 63 percent greater if training costs remained constant between 1980 and the early 2000s.

Keywords: Training, Technological Change, College Wage Premium, Education, Technology
JEL Codes: E24, I24, J24, J31, M53, O33

*I am grateful to my advisors, Bettina Brueggemann, Pau Pujolas, Zachary Mahone, and Gajendran Raveendranathan, for their feedback throughout the course of this project. I also thank José Víctor Ríos Rull, Alok Johri, Oliver Loertscher, Andrew Leal, participants at the CEA 2021 and 2022 annual meetings, and participants at the 2022 PAELLA Workshop for their helpful comments. I acknowledge support from the Productivity Partnership and the Social Sciences and Humanities Research Council for Doctoral Fellowship 752-2022-1241. All remaining errors are mine.

†Department of Economics, McMaster University, Hamilton, ON, Canada, L8S 4M4. Email: palmet8@mcmaster.ca. Website: www.thomasmarkpalmer.com.

1 Introduction

The proliferation of new information-and-communications technologies (ICT) beginning in the 1980s coincided with a rise in both the relative supply and price of post-secondary educated labor. The starkest changes occurred in the United States, where the college wage premium more-than-doubled and the share of post-secondary educated workers increased by 91 percent over this period. Other OECD countries experienced a similar fate, although typically to a lesser degree. For example, in Canada, the college wage premium increased by 22 percent between 1980 and 2000[1], while the share of post-secondary educated (henceforth, high-skill) workers increased by 67 percent. In this paper, I document a new channel connecting technological change to the college wage premium: employer-provided training. In particular, I empirically document that new technologies often require additional training, which itself generates an earnings premium independent of education. I then quantitatively show that increased training participation among low-skill workers has produced a dampening effect on the college wage premium.

Existing studies of the college wage premium, including Bound and Johnson (1992), Katz and Murphy (1992), and Krusell et al. (2000), often highlight the effect of new technologies on the productivity of high-skill workers and, hence, how skill-biased technological change (SBTC) raises the college wage premium through the response of labor demand. The main empirical contribution of this paper is to document a related link between technology and training provision at the workplace-level. Specifically, I use unique matched employer-employee survey data from Canada to establish three facts. First, I show that technology-intensive (henceforth, high-technology) firms are relatively more productive, employ more high-skill workers, and provide more training than low-technology firms. That is, there exists a positive relationship between work-related training and technological intensity in the workplace. Second, I demonstrate that training generates a significant earnings premium, conditional on various worker and firm characteristics. Third, I show that the difference in training participation rates between high- and low-skill workers is smaller within high-technology firms. As technological change intensifies, the relative training participation rate among low-skill workers accelerates which, in turn, raises the average wage among this group and reduces the college wage premium. I provide further evidence of this mechanism by estimating a college wage premium separately for trained and untrained workers within high-technology firms and show that the premium is indeed smaller for the trained group.

Together, Facts 1 and 2 suggest that high-skill workers disproportionately benefit from technological change: not only does the return to their formal credentials increase, but they are also more likely to work for high-technology employers and reap the benefits of training over their working

[1] See, for example, Krueger et al. (2010). In related work, Boudarbat, Lemieux, and Riddell (2010) document an increase of 25 percent for the university-to-high-school graduate premium among males between 1980 and 2005 using Canadian Census data. On the other hand, Kryvtsov and Ueberfeldt (2009) find no change in the premium between 1980 and 2000 when comparing males with at least a bachelor's degree to males without any post-secondary education using Canadian Survey of Consumer Finances data.

career. When combined with Fact 3, however, the opposite story begins to emerge. Therefore, to better understand the aggregate implications of these facts, I build a quantitative model of the labor market, which embeds endogenous investments in education (Flinn and Mullins, 2015; Shephard and Sidibe, 2019) for workers, and in training (Moen and Rosén, 2004; Flinn, Gemici, and Laufer, 2017) and technologies (Shi, 2002) for firms, into a general equilibrium directed search model. Assuming that the benefits of training are fully match-specific, search frictions allow workers to extract some of the rents from training.[2] To ensure interior solutions for the shares of high-skill, trained, and high-technology employees as in the data, I allow the cost functions for technology investment and training provision to depend on the education level of a worker. With directed search, solving the model remains tractable: given a set of taxes to finance the economy's unemployment insurance program, the worker and firm problems can be solved independently of the distribution of individuals across states.[3] The model generates an exact decomposition of the college wage premium into three channels: (1) the relative supply high-skill workers, (2) the relative complementarity between high-skill labor and technology, and (3) the training participation gap between high- and low-skill workers.

I calibrate the model starting with the final steady state and work backwards in time. I pool repeated cross-sections from the Canadian micro-data between 1999 and 2005 to obtain a set of 11 moments to exactly identify 11 parameters in the final steady state. To identify and estimate the parameters governing the initial steady state, I impose an additional restriction. Specifically, I sort the set of 11 parameters into a group of 7 parameters, which are held fixed across the steady states, and a group of 4 parameters, which vary across the steady states. The first group of parameters governs the productivity of education and training, the cost of posting vacancies, and the setup cost associated with the high technology. The second group of parameters governs the complementarity between technology and high-skill labor, the costs of training, and the cost of higher education; their values are chosen to match the increase in the college wage premium between 1980 and 2000, the share of training participants in 1980, and the share high-skill workers in 1980, respectively.

I use the calibrated model as a laboratory to decompose the college wage premium and, ultimately, measure the effect of training on the observed rise in the college wage premium for Canada between 1980 and the early 2000s. Consistent with existing quantitative work, I find that technological change is the primary driver of the college wage premium and explains 60 percent of its absolute variation over this period. Increased training participation explains a meaningful 28 percent of the absolute variation and, in fact, dampens the premium. In particular, I show that, if training costs were held fixed between the 1980 and the early 2000s, the increase in the college wage premium would have been 63 percent greater.[4]

[2] Put differently, in a frictionless environment with fully match-specific training, workers would not receive any premium from training, which is inconsistent with the data.

[3] See, also, Menzio and Shi (2010, 2011).

[4] In terms of its level, the college wage premium under fixed training costs would be 16 percent larger than its 1980 level or 6 percent larger than its 2000 level.

The remainder of the paper is organized as follows. In Section 2, I discuss the related literature on training, technological change, and the college wage premium. In Section 3, I describe the data in detail and present the motivating facts. In Section 4, I formally describe the model and define the equilibrium concept. In Section 5, I discuss the calibration. In Section 6, I present the results of the decomposition exercise and discuss its implications. Section 7 concludes the paper.

2 Related Literature

Starting with the seminal contributions of Bound and Johnson (1992), Katz and Murphy (1992), Autor, Katz, and Krueger (1998), and Krusell et al. (2000), an extensive literature has explained the rise in the college wage premium experienced by the U.S. between the 1960s and early 2000s as the result of skill-biased technological change (SBTC). The unifying theme of this literature is that the introduction of new production technologies—notably, computers—has disproportionately benefited high-skill workers because of the existence of capital-skill complementarities in production. As a result, a fall in the relative price of capital goods increases the relative demand for high-skill labor and exacerbates the college wage premium. Katz and Murphy (1992) formally demonstrate this mechanism by applying a competitive model of the labor market to data from the Current Population Survey for the period 1963-1987. Krusell et al. (2000) enrich the analysis by developing a model that links technological change to observables and use it to decompose the college wage premium over a longer time horizon. Ultimately, they find that the combination of cheaper capital goods and capital-skill complementarity accounts for approximately two-thirds of the growth in the college wage premium observed in the U.S. from 1963 to 1992.[5]

Clearly, however, there are many factors beyond technological change that also affect the college wage premium.[6] For example, Walker and Zhu (2008), Velden and Bijlsma (2016), and Matsuda (2020) conduct a more thorough analysis of the college enrollment decision and document the importance of accounting for large changes in the relative supply of high-skill labor when analyzing changes in the college wage premium. Parro (2013), Dix-Carneiro and Kovak (2015), and Burstein and Vogel (2017) demonstrate that reductions in trade costs exacerbate the college wage premium at both an aggregate and a local level. Açıkgöz and Kaymak (2014) and Zentler-Munro (2021) study the role of bargaining power by drawing attention to the different rates of deunionization faced by high- and low-skill workers over time. Finally, He (2012) studies the impact of large-scale changes in the age structure of the economy—specifically, the baby boom and baby bust—on college enrollment and the college wage premium. While these alternative and complementary

[5]Several papers have also applied similar frameworks to analyze trends in college wage premia in other countries, including Canada (Burbidge, Magee, and Robb, 2002; Boudarbat, Lemieux, and Riddell, 2010), the U.K. (Blundell, Green, and Jin, 2022), Indonesia (Amiti and Cameron, 2012), Japan (Lise et al., 2014; Takahashi and Yamada, 2022), Germany (Glitz and Wissmann, 2021), and many others.

[6]See, also, Card and DiNardo (2002) for a more complete discussion of some prominent alternatives to the SBTC hypothesis.

explanations certainly offer useful insights, the studies mentioned above generally still require a strong role for technology—and, in particular, capital-skill complementarities in production—to generate an empirically-consistent rise in the college wage premium. For this reason, I ultimately focus on technological change as the main driver of the college wage premium.

Within this literature, the most closely related papers are Lindner et al. (2022) and Doepke and Gaetani (2020). Lindner et al. (2022) study the impact of firm-level technological change on skill demand and aggregate inequality using a quantitative model of an imperfectly competitive labor market. In their framework, firms' wage policies internalize the fact that higher wages attract more workers. Therefore, in response to skill-biased technological change, firms increase the relative wage of high-skill workers, which generates a corresponding increase in both the firm's share of high-skill workers and the aggregate college wage premium. In my framework, skill-biased technological change similarly increases the share of high-skill workers and the college wage premium but for *two* reasons. First, SBTC directly increases the productivity of, and return to, being a high-skill worker. New firms respond to technological change by posting vacancies to attract high-skill workers, while more newborn individuals respond by enrolling in college at a greater rate. Second, the additional productivity gain from SBTC experienced by all firms encourages more of them to provide training. Workers in high-technology firms especially benefit, as they receive both a direct increase in productivity and a indirect increase in the probability of receiving training. Doepke and Gaetani (2020) study the differences in college wage premia between the U.S. and Germany through the lens of a competitive model of the labor market in which firms and workers make match-specific investments in skill accumulation. Because the incentive to invest in skills is increasing in the expected duration of the match, they argue that stricter employment protection laws in Germany have particularly benefited low-skill workers and have helped to moderate the German college wage premium over time. Contrary to Doepke and Gaetani (2020), I do not consider differences in employment protection. Instead, I show that technology itself generates a greater incentive to provide work-related training. The effect is most pronounced for low-skill workers, thus attenuating the college wage premium over time.

There is also a large empirical literature devoted to estimating the causal effect of work-related training on earnings. Studies in this literature generally find that training generates large and persistent returns for participants. For example, Blundell, Dearden, and Meghir (1996) apply a quasi-difference specification to a subset of individuals from the British National Child Development Survey and estimate a 3.6 percent wage increase to male participants of employer-provided on-the-job training courses and a 6.6 percent wage increase for employer-provided off-the-job training courses. Parent (1999) estimates a series of OLS and IV regressions using data from the National Longitudinal Survey of Youth (NLSY) and finds a wage effect ranging from 12 to 17 percent for on-the-job training and 7.5 to 14 percent for off-the-job training provided by an individual's current employer. Leuven and Oosterbeek (2008) consider a novel identification strategy by comparing training participants to non-participants who initially wanted to participate but were unable to for exogenous reasons. Under this environment, the authors estimate a near-zero return to training participation, suggesting that there may be some selection in terms of who does and does not receive training. In my empirical analysis of Canadian micro-data, I show that training participation is indeed skewed toward specific worker and firm types: those who are high-skill and employed at high-technology firms.

Finally, a few papers have attempted to link technology adoption, training, and labor market outcomes in their empirical work. Bartel and Sicherman (1998) evaluate the effects of technological change on individual training participation by combining the NLSY with cross-sectional industry-level measures of computer investment, total factor productivity, and R&D intensity. Ultimately, they find that workers are more likely to receive training as technological progress intensifies within the industry, and that training participation rates are increasing in education. Relative to Bartel and Sicherman (1998), the data I use allows me to identify technological change at a finer level, namely, the firm. As a result, I show that the positive association between technology and training carries over to the individual match-level and extends more generally to industries beyond manufacturing. In more recent work, Bresnahan, Brynjolfsson, and Hitt (2002), Black and Lynch (2004), and Boothby, Dufour, and Tang (2010) use firm-level survey data to study how firms' decisions to adopt new technologies affects productivity and innovation through its interaction with workplace organization and training provision. In contrast to this set of papers, I link information on technology use at the workplace level to individual employee characteristics. Accordingly, I not only provide further evidence that high-technology firms offer more training in absolute terms, but also show that the increased likelihood of receiving training in such firms is relatively greater for low-skill workers.

3 Data

In this section, I document new facts related to technology, training, and the college wage premium based on survey micro-data from Canada. Together, the facts demonstrate the key mechanism underlying the quantitative model of Section 4.

3.1 Workplace and Employee Survey (WES)

The empirical analysis is based on data from the Workplace and Employee Survey (WES). The WES is a matched employer-employee survey data set from Canada, which covers approximately 20,000 employees spread across 6,000 workplaces at an annual frequency from 1999 to 2006.[7] I focus, in particular, on the cross-sectional workplace and employee samples from 1999, 2001, 2003, and 2005.[8]

[7] A "workplace" in this context means "establishment."

[8] Because sampled workers are only followed for two years at most, the survey-weighted statistics computed on the even-year samples of employees reflect the population statistics for the preceding (odd) survey year (minus attrition). Further details about the data and sampling design are contained in Appendix A.

I also restrict attention to workers aged 25 to 64 years old to limit variation in hours, employment status, and earnings arising from full- or part-time enrollment in education and retirement.

The WES data provides two main benefits for studying the link between technological change and the college wage premium. First, the data contain rich survey information on technological intensity and adoption at the workplace-level. Existing studies of technology adoption and wage inequality, such as Autor, Katz, and Krueger (1998) and Kristal (2020), are typically restricted to industry- or occupation-level analyses. Instead, I exploit the workplace-level variation in the WES to examine the effects of technology at a more granular level. Second, the data contain information on training participation and provision, thereby overcoming a major challenge in the existing empirical literature on human capital accumulation. Importantly, the ability to link employees to their employers is crucial to identify the responsiveness of training to technology, which is a key mechanism driving changes in the college wage premium in the model.

3.2 Definitions of Training, Education, and Technological Intensity

3.2.1 Training Participation and Provision

The WES collects information on two main types of employer-provided training: formal classroom and informal on-the-job training. For each mode of training, both employees and employers are asked about training duration, intensity[9], subject matter, and funding sources. Since both training types are identified as *employer-provided*, I define training participants as workers who have received *either* classroom *or* on-the-job training over the past survey year and training firms as workplaces that have provided *either* classroom *or* on-the-job training to at least one employee over the past survey year.

Under this classification, approximately 50 to 60 percent of workers are identified as training participants and 40 to 50 percent of workplaces are identified as training providers each year, on average.

3.2.2 High-Skill and Low-Skill Workers

In every odd survey year, employees are asked the following series of questions:

1. *Did you graduate high-school?*
2. *Have you received any other education?*
3. *What was that education?*

I define *high-skill* workers as employees that have graduated high school (answered "yes" to question 1) and have received additional education (answered "yes" to question 2). The set of high-skill employees thus includes college and university graduates, as well as employees with post-secondary

[9] For example, a measure of how many courses were taken.

diplomas; trade, vocational, or industry certificates; and, post-graduate or professional degrees. On the other hand, I define *low-skill* workers as employees that have either not graduated high-school (answered "no" to question 1) or have graduated high-school without any additional education (answered "yes" to question 1 and "no" to question 2).

On average, high-skill workers account for approximately two-thirds of the labor force over the sample period. Roughly 25 to 30 percent of the overall sample are university-educated (Bachelor's, Master's, Ph.D., M.D., etc.); 30 to 35 percent are college-educated (college degrees, trade or vocational school, industry certified, etc.); and the remainder are high-school graduates and dropouts.

3.2.3 High-Tech and Low-Tech Firms

In each survey year, the WES questionnaire asks employers the following question:

> *"At this location, how many employees currently use computers as part of their normal working duties? By computers, we mean a microcomputer, personal computer, minicomputer, mainframe computer or laptop that can be programmed to perform a variety of operations."*

Using the reported answers to this question and the workplaces' total number of employees, I construct a variable that identifies the *share* of a workplace's employees who use computers regularly on the job. Formally, for each workplace j in year t, I compute:

$$\text{ShareCPU}_{j,t} = \frac{\text{Number of Computer Users}_{j,t}}{\text{Total Number of Employees}_{j,t}}. \tag{1}$$

I use the estimated shares to classify employers as high- or low-technology. I define *high-technology* firms as workplaces with at least 50 percent of employees using computers as part of their normal working duties: ShareCPU ≥ 0.50. Low-technology firms comprise the remaining firms, that is, workplaces with strictly less than 50 percent of employees using computers: ShareCPU < 0.50.[10]

3.3 Stylized Facts

Having defined the main variables of interest, I now establish empirically how technological change affects the college wage premium through the response of training. I report these findings as a series of three stylized facts.

Fact 1: High-technology firms are more productive, more likely to provide training, and more likely to employ high-skill workers than low-technology firms.

[10] Autor, Katz, and Krueger (1998) use a similar strategy to identify high-technology industries. However, rather than using a continuous measure of technological intensity, I adopt a binary classification for consistency with my quantitative model.

To understand the distinguishing characteristics of high- and low-technology firms, I evaluate their differences along three dimensions: (1) productivity, (2) training, and (3) employment structure.

Table 1 presents the results from estimating a workplace-level regression of (log) revenue productivity on a high-technology indicator; a vector of time-varying workplace-level control variables, including industry and firm size; and year fixed-effects.[11] High-technology firms are, on average, approximately 58 percent more productive than technology firms. This finding is consistent with existing models of technology investment and firm dynamics, such as Shi (2002), in which a firm invests in the high-technology whenever it meets or exceeds a threshold level of idiosyncratic productivity.

Table 1: Relative Productivity of High-Tech Firms

Dependent Variable: Log[Revenue Productivity]	
High-Tech	0.4538***
	(0.0306)
N (Unweighted)	22,392
N (Weighted)	2,627,197

Note: Standard errors in parentheses. Standard errors are bootstrapped using the workplace bootstrap weights provided by Statistics Canada and 100 replications. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.10$

To evaluate differences in training incidence between high- and low-technology firms, I estimate an employee-level discrete choice regression of training participation on a high-technology indicator. The top panel of Table 2 summarizes the estimated odds ratios, while the bottom panel reports the average marginal effects. Column (1) reports the results of a regression controlling only for the worker's skill level, while Column (2) includes a richer set of time-varying individual-level control variables (still including skill level), a set of time-varying workplace-level control variables, and year fixed-effects. The results here imply that high-technology employees are, on average, 10.4 percent more likely to receive training, conditional on education, and 3.4 percent more likely to participate in training, after controlling for a host of additional observables.

[11] The specifications for all regressions underlying Fact 1 are stated formally in Appendix B.

Table 2: Probability of Training Participation

Dependent Variable: Training Indicator	**(1)**	**(2)**
Odds Ratios		
High-tech	0.4311***	0.1573**
	(0.0330)	(0.0617)
High-skill	0.6847***	0.2895***
	(0.0335)	(0.0385)
Average Marginal Effects		
High-tech	0.1036***	0.0337**
	(0.0184)	(0.0133)
High-skill	0.1673***	0.0626***
	(0.0184)	(0.0084)
Additional Controls?	–	✓
N (Unweighted)	81,622	65,814
N (Weighted)	41,405,006	31,047,424

Note: Standard errors in parentheses. Standard errors are bootstrapped using the workplace bootstrap weights provided by Statistics Canada and 100 replications. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.10$

Finally, I assess whether the definitions of high- and low-technology firms I use are consistent with the existing literature on technology-skill complementarities by estimating an employee-level logistic regression of the high-technology indicator on an indicator for the worker's skill level. As before, the top panel of Table 3 reports the estimated odds ratios and the bottom panel reports the average marginal effects. Column (1) controls only for training participation, while Column (2) controls for additional workplace- and worker-level variables including occupation, age, experience, gender, CBA coverage, immigration status, industry, firm size, and year. The estimated marginal effects imply a 19 percent greater likelihood for high-skill workers to be employed in a high-technology firm, conditional on receipt of training, and a 2.3 percent greater likelihood, conditional on a series of additional covariates.

Table 3: Probability of High-Technology Employment

Dependent Variable: High-tech Indicator	**(1)**	**(2)**
Odds Ratios		
High-skill	0.7881***	0.1448***
	(0.0404)	(0.0499)
Trained	0.4311***	0.1447**
	(0.0330)	(0.0621)
Average Marginal Effects		
High-skill	0.1916***	0.0233***
	(0.0095)	(0.0080)
Trained	0.1039***	0.0232**
	(0.0079)	(0.0100)
Additional Controls?	–	✓
N (Unweighted)	81,622	65,532
N (Weighted)	41,405,006	31,047,424

Note: Standard errors in parentheses. Standard errors are bootstrapped using the workplace bootstrap weights provided by Statistics Canada and 100 replications. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.10$

Fact 2: Training participation generates a significant premium on hourly earnings.

While Fact 1 establishes a positive relationship between technology and training, it remains to show that training by itself has an effect on the college wage premium separate from technology. To address this issue, I perform a series of Mincer (1958) style regressions to quantify the impact of employer-provided training on hourly earnings at the worker-level, both conditionally and unconditionally. For a worker i employed by firm j in year t, the baseline specification is given by:

$$\begin{aligned} \ln(\text{Earnings}_{i,t}) = \beta_0 + \beta_1 \text{Train}_{i,t} + \beta_2 \text{HighSkill}_{i,t} + \beta_3 \text{HighTech}_{j,t} + \beta_4 (\text{HighSkill}_{i,t} \times \text{HighTech}_{j,t}) \\ + \theta_t + \delta X_{i,t} + \xi Z_{j,t} + \varepsilon_{i,t} \end{aligned} \tag{2}$$

where hourly earnings are expressed in constant 1999 dollars; the vector $X_{i,t}$ of individual-level control variables includes occupation, usual weekly hours, tenure, experience, age, an indicator for non-employer-sponsored career-related training, CBA coverage, gender, immigrant status, and indicators for whether the worker uses computers or other types of information technologies in the workplace; and the vector $Z_{j,t}$ of time-varying firm-level controls includes firm size, productivity, and industry.

The coefficient of interest is β_1 and identifies the impact of training participation on (log) hourly earnings. Importantly, I also control for the effects of workers' skills (captured by β_2), the

Table 4: Impact of Training Participation on Earnings

Dependent Variable: Log[Hourly Earnings]	**(1)**	**(2)**
Trained	0.1407***	0.0292***
	(0.0119)	(0.0079)
High-skill	0.1804***	0.0641***
	(0.0110)	(0.0090)
High-tech	0.1310***	0.0120
	(0.0171)	(0.0125)
High-skill × High-tech	0.063***	0.0737***
	(0.0203)	(0.0166)
Additional Controls?	–	✓
N (Unweighted)	81,622	65,532
N (Weighted)	41,405,006	30,899,192

Note: Standard errors in parentheses. Standard errors are bootstrapped using the employee bootstrap weights provided by Statistics Canada and 100 replications. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.10$

technology of worker i's employer (captured by β_3), and an interaction between worker skills and employer technology (captured by β_4) to control for potential earnings effects from technology-skill complementarities. Table 4 reports the results from performing this regression after accounting for the complex survey design of the WES.

Column (1) on the left reports the results from the regression omitting all control variables except for the ones specified—that is, without θ_t, $X_{i,t}$, and $Z_{j,t}$. Column (2) on the right reports the results from the regression with additional controls. In both cases, training participation is shown to have an economically and statistically significant impact on individual earnings. For the baseline case—conditioning only on education and employer type—training participation is associated with a roughly 15 percent increase in hourly earnings, while in the preferred specification with additional controls, training participation is associated with a 3 percent increase in hourly earnings. The positive wage effect of training is robust to various alternative sets of control variables and clustering standard errors by workplace.

Fact 3: Technology-induced training dampens the college wage premium.

The main limitation of the WES micro-data is the short time horizon that it covers: it is simply not possible to obtain a direct measure of how training has evolved for high- and low-skill workers over time. However, it *is* possible to recover an indirect estimate of the evolution by exploiting the properties of technological change. In particular, if a relatively larger share of firms in 1980 were low-technology, then one way to indirectly assess the evolution of training participation by skill level is to measure how much the rate of training participation increases for each group when comparing low-technology to high-technology employees in the cross-section. Performing this exercise yields the results reported in Table 5.

Columns 2 and 3 list the share of training participants within low- and high-technology firms,

Table 5: Training Participation Rates by Education and Technology

	Training Participation Rates (%)		
Skill-Group	**Low-Tech**	**High-Tech**	**Change**
Low-Skill	37.48	50.62	35.04
High-Skill	55.95	65.07	16.30

Note: The column "Change" reports the difference in training participation rates (in percent) between low- and high-technology firms, conditional on skill group.

respectively, conditional on a worker's skill level. Column 4 reports the relative increase in training participation for a worker of a given skill level when moving from a low- to a high-technology firm. As illustrated, the increased likelihood of receiving training in high-technology firms is much more pronounced for low-skill workers: low-skill workers experience a 35 percent increase in the probability of receiving training, while high-skill workers experience a 16 percent increase. Put differently, the relative training participation rate for low-skill workers is 67 percent in low-technology firms and 78 percent in high-technology firms. A broader implication of this result is that, conditional on being employed in a high-technology firm, the college wage premium is smaller among the subset of trained workers. This is precisely the result shown in Figure 1.

4 Model

To better understand the aggregate implications of the stylized facts presented in Section 3, I now develop a quantitative model of the labor market, which connects human capital investments in education and training to an aggregate process of skill-biased technological change in the spirit of Shi (2002) and Salgado (2020). The model features two levels of education and two levels of production technologies, as in the data. The shares of worker and firm types are determined endogenously in equilibrium. Workers decide their level of education prior to entering the labor market by solving a trade-off between the financial and non-pecuniary cost of enrollment and the expected return, which internalizes the probability of finding employment and receiving training. On the other side of the market, firms choose their production technologies at the time of posting vacancies by solving a trade-off between the higher setup cost associated with the high-technology and the expected return, which internalizes the firm's optimal training and wage policies conditional on matching.

4.1 Environment

Time is discrete and continues forever, $t = 0, 1, 2, \ldots$. Three types of agents populate the economy: workers, firms, and a government. The mass of workers is normalized to 1 and the mass of firms is endogenously determined through free entry into the labor market. Workers are ex-ante heterogeneous in innate ability a, risk neutral, and discount future income with factor $\beta \in (0, 1)$.

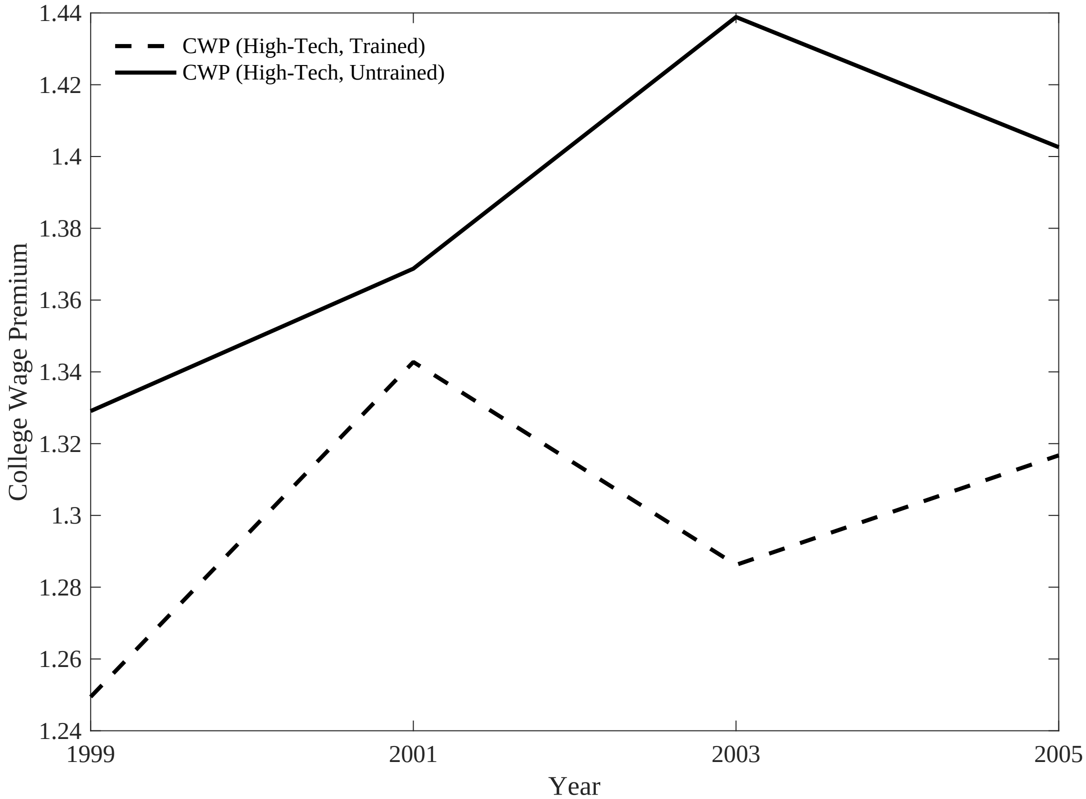

Figure 1: College Wage Premium by Technology and Training

I assume that innate ability a takes on one of m_a possible values, that is, $a \in \{a_1, a_2, \ldots, a_{m_a}\}$. Firms are ex-ante heterogeneous in productivity z, risk neutral, and share the same discount factor β. I assume that firm productivity z takes on one of m_z possible values, that is, $z \in \{z_1, \ldots, z_{m_z}\}$. The government sets lump-sum taxes τ on labor income to finance the economy's unemployment insurance program.

As workers choose their level of education $s \in \{0, 1\}$ prior to entering the labor market, it is convenient to envision each period unfolding in two stages: an education stage and a labor market stage. In the education stage, individuals decide—conditional on their innate ability a—whether to invest in post-secondary education by paying the cost of enrollment $c_s - \zeta_s > 0$ and enter the labor market as a type-$(a, 1)$ high-skill worker, or to forgo additional education and enter the labor market as a type-$(a, 0)$ low-skill worker. Once a worker enters the labor market, their education level s is fixed for the remainder of their life. In the labor market stage, workers simply shift between employment and unemployment depending on the shocks they experience over their working life.

The labor market is composed of a continuum of submarkets indexed by the set (a, s, z, k, x, ω), consisting of worker innate ability a, worker education $s \in \{0, 1\}$, firm productivity z, firm technology $k \in \{0, 1\}$[12], firm training provision $x \in \{0, 1\}$, and the piece-rate $\omega \in [0, 1]$ posted by the firm. The piece-rate represents the fraction of output promised to the worker each period. For convenience, let $\mathbf{s}_w = (a, s)$ represent the vector of worker types and $\mathbf{s}_f = (z, k, x)$ represent the vector of firm types. There is no uncertainty. Search is directed in the sense that an unemployed worker of type $\mathbf{s}_w$ observes the firm types $\mathbf{s}_f$ and piece-rates ω posted within each submarket prior to making their application decision. The assumptions regarding the information environment imply that, when searching, an unemployed worker knows whether the firm has invested in the high-technology and whether they will receive training when computing their expected value from employment; and there is no incentive for either party to renegotiate contracts ex-post once a match is formed.

Matches within each submarket are governed by a constant-returns-to-scale matching function $m(u, v)$, where u represents the mass of unemployed searchers and v represents the associated mass of vacancies in a given submarket. Let $\theta \equiv v/u$ denote the ratio of vacancies to unemployment, or tightness, of a submarket. Then, for a submarket with tightness $\theta(\mathbf{s}_w, \mathbf{s}_f, \omega)$, the probability that a worker successfully finds a job is given by:

$$p\left(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega)\right) = \frac{m\left(u(\mathbf{s}_w, \mathbf{s}_f, \omega), v(\mathbf{s}_w, \mathbf{s}_f, \omega)\right)}{u(\mathbf{s}_w, \mathbf{s}_f, \omega)} \in (0, 1). \tag{3}$$

On the other hand, the probability that a vacant job successfully finds a worker is given by:

$$q\left(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega)\right) = \frac{m\left(u(\mathbf{s}_w, \mathbf{s}_f, \omega), v(\mathbf{s}_w, \mathbf{s}_f, \omega)\right)}{v(\mathbf{s}_w, \mathbf{s}_f, \omega)} \in (0, 1). \tag{4}$$

Matches only end for exogenous reasons. With probability $\eta_s \in (0, 1)$, an existing match is exoge-

[12]Here, $k = 1$ represents the high-technology while $k = 0$ represents the low-technology.

nously separated each period; and, with probability $\delta \in (0,1)$, an existing worker exogenously leaves the economy each period. When a match receives a separation shock η_s, the job is destroyed and the worker transitions to unemployment. When a worker receives an exit shock δ, the job is destroyed and the worker is replaced by a newborn worker of the same type, who enters the education stage next period.[13]

Timing within a period occurs as follows. First, a mass δ of newborn workers make their education decisions. In the following period, these workers will enter the labor market as a type-$\mathbf{s}_w$ unemployed worker. Second, idiosyncratic shocks are realized. Existing matches are exogenously separated with probability η_s and existing labor market participants exogenously exit the market with probability δ. Third, potential entrant firms pay cost κ to draw productivity z and, conditional on the realization of productivity, decide whether to invest in the high-technology. Given its productivity and technology, the firm chooses a contract consisting of a piece-rate ω and training policy x to post in a submarket to attract type-$\mathbf{s}_w$ unemployed workers. On the other side of the market, unemployed type-$\mathbf{s}_w$ workers direct their search to the submarket that maximizes their lifetime value. Fourth, matching occurs. Unfilled vacancies and unemployed job seekers are matched in each submarket according to the matching function $m(u, v)$ described above. Fifth, production occurs. Firms execute their contractual obligations by paying the training cost $c_x(s)$ and produce output according to the production function $y(\mathbf{s}_w, \mathbf{s}_f)$. Finally, payments are made. Firms earn operating profits $(1-\omega)y(\mathbf{s}_w, \mathbf{s}_f)$, employed workers earn net income $\omega y(\mathbf{s}_w, \mathbf{s}_f) - \tau$, and unemployed workers earn unemployment insurance b_s.

4.2 Firms

4.2.1 Production

There is a large mass λ of firms each period, which is determined by free entry. Each firm employs at most one worker. When matched with a worker of type-$\mathbf{s}_w$, a firm of type-$\mathbf{s}_f$ produces output $y(\mathbf{s}_w, \mathbf{s}_f)$ according to the following production technology:

$$y(\mathbf{s}_w, \mathbf{s}_f) = \varepsilon_x(s)^x \varepsilon_k(s)^k zah_s, \tag{5}$$

where $\varepsilon_x(s)$ and $\varepsilon_k(s)$ capture, in a reduced-form way, the productivity gains earned through training and technological investments by the firm; and, h_s represents the productivity gain earned through investments in education by the worker. I assume that $\varepsilon_x(s) > 1$ and $\varepsilon_k(s) > 1$ for all s, $h_0 = 1$, and $h_1 > 1$.

[13]Hence, with a mass of workers normalized to 1, δ represents both the probability of exiting and the mass of entering workers each period.

4.2.2 Firm Value Function

Consider a match of type $(\mathbf{s}_w, \mathbf{s}_f, \omega)$ and let $\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega)$ denote the continuation profit for the firm—that is, the profit accrued to the firm in every period after paying the training cost in period 1. Then, $\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega)$ solves the following recursive equation:

$$\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega) = (1-\omega)y(\mathbf{s}_w, \mathbf{s}_f) + \beta(1-\delta)(1-\eta_s)\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega). \tag{6}$$

Let $J_F(\mathbf{s}_w, \mathbf{s}_f, \omega)$ denote the lifetime value for the same firm. The only difference between $\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega)$ and $J_F(\mathbf{s}_w, \mathbf{s}_f, \omega)$ is that the latter includes a one-time training cost, conditional on providing training. That is, $J_F(\mathbf{s}_w, \mathbf{s}_f, \omega)$ solves:

$$J_F(\mathbf{s}_w, \mathbf{s}_f, \omega) = (1-\omega)y(\mathbf{s}_w, \mathbf{s}_f) - c_x(s)x + \beta(1-\delta)(1-\eta_s)\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega), \tag{7}$$

where $c_x(s) > 0$ is the lump-sum cost of training. In the current period, a firm earns revenue from output $y(\mathbf{s}_w, \mathbf{s}_f)$ net of the wage bill $\omega y(\mathbf{s}_w, \mathbf{s}_f)$ and training cost $c_x(s)$, conditional on training $(x = 1)$. In the following period, the match survives with probability $(1-\delta)(1-\eta_s)$, in which case the firm earns continuation value $\Pi(\mathbf{s}_w, \mathbf{s}_f, \omega)$. Hence, when a firm commits to providing training, $x = 1$, it must pay for training *once* even though the return to training—captured in $\varepsilon_x(s)$—accrues to the match for as long as it survives.

4.2.3 Free Entry and the Zero Profit Condition

Each period, firms enter submarkets until the value of a vacancy in each submarket is driven to zero. By paying cost κ, a potential entrant draws a level of productivity z and decides whether to invest in the high-technology, $k \in \{0, 1\}$. Conditional on the choice of technology k and its productivity z, the firm must also decide the terms of the contract (x, ω) to post along with its vacancy. Hence, the zero profit condition for submarkets is given by:

$$\kappa + c_k(s)k = q(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))J_F(\mathbf{s}_w, \mathbf{s}_f, \omega), \tag{8}$$

where $q(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))$ is the job-filling probability in submarket $(\mathbf{s}_w, \mathbf{s}_f, \omega)$. Therefore, firms who invest in the high-technology pay cost $\kappa + c_k(s)$ to post a vacancy, while firms who maintain the low-technology only pay κ. The solution to this problem yields tightness $\theta(\mathbf{s}_w, \mathbf{s}_f, \omega)$ for each submarket and job finding probabilities $p(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))$, which are required to solve the worker's problem.

4.3 Workers

The mass of workers is normalized to 1. Over their lifetime, workers occupy three possible states: employed at a firm of type-$\mathbf{s}_f$ and receiving piece-rate ω; unemployed and earning unemployment insurance b_s; or out of the labor force. Note that workers who are out of the labor force only consist of newborn workers who have yet to make their education decision—there is no explicit labor force participation decision.

4.3.1 Employed Value Function

Consider first an employed worker of type-$\mathbf{s}_w$, who is working at a firm of type-$\mathbf{s}_f$ that provides a piece-rate of ω each period. Let $V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$ denote the lifetime value for this worker. Each period, the worker simply consumes their earnings, while facing exogenous probabilities $\eta_s \in (0, 1)$ and $\delta \in (0, 1)$ of getting hit by a job separation or exit shock, respectively. In the case of receiving a separation shock η_s, the worker transitions to unemployment and receives continuation value $V_U(\mathbf{s}_w)$; in the case of receiving an exit shock, the worker receives a value of 0 and is replaced by a newborn worker of the same ability a in the following period. Conditional on surviving both shocks, however, the worker retains the value $V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$ from employment into the next period. Accordingly, the value function $V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$ solves the following recursive equation:

$$V_E(\mathbf{s}_w, \mathbf{s}_f, \omega) = \omega y(\mathbf{s}_w, \mathbf{s}_f) - \tau + \beta(1-\delta)\Big[\eta_s V_U(\mathbf{s}_w) + (1-\eta_s)V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)\Big], \tag{9}$$

where $\tau \geq 0$ is a lump-sum tax on labor earnings collected by the government.

4.3.2 Unemployed Value Function

Each period, an unemployed worker of type-$\mathbf{s}_w$ decides which submarket $(\mathbf{s}_w, \mathbf{s}_f, \omega)$ to search in. With probability $p(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))$, the worker successfully finds a job offering piece-rate ω and receives value $V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$ in the next period conditional on survival; otherwise, the worker remains unemployed and receives value $V_U(\mathbf{s}_w)$. Hence, $V_U(\mathbf{s}_w)$ solves:

$$V_U(\mathbf{s}_w) = \max_{(\mathbf{s}_f, \omega)} \; b_s + \beta(1-\delta)\Big[p(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))V_E(\mathbf{s}_w, \mathbf{s}_f, \omega) + (1 - p(\theta(\mathbf{s}_w, \mathbf{s}_f, \omega))V_U(\mathbf{s}_w)\Big]. \tag{10}$$

4.3.3 Education Decision

Newborn workers with innate ability a make their education decisions at the start of each period. This decision involves paying a cost $c_s - \zeta_s > 0$ to enroll in education level $s = 1$ and earn human capital h_s, where I assume that $h_0 = 1$ and $h_1 > 1$. To understand this formulation of the cost function, it is convenient to think of c_s capturing the financial costs of enrollment and ζ_s representing unmodeled preferences for higher education, which may include various non-monetary costs associated with enrollment.

I assume that the stochastic component of the cost function, ζ_s, follows a Type-I Extreme Value distribution with zero location parameter and scale parameter $\chi_s > 0$. Following this decision, a worker enters the labor market as a type-$\mathbf{s}_w$ unemployed worker in the following period. Accordingly, the optimal education policy for a worker with innate ability a, before the realization of the shock

ζ_s, solves:

$$s(a) = \underset{s\in\{0,1\}}{\arg\max}\ \mathbb{E}_{\zeta_s}\left\{s\left[\beta(1-\delta)V_U(a,1) - c_s + \zeta_s\right] + (1-s)\beta(1-\delta)V_U(a,0)\right\} \tag{11}$$

The introduction of Type-I Extreme Value shocks into the education decision implies that a worker's education policy is a probability.

4.4 Government

The government levies lump-sum taxes on labor income τ to finance unemployment insurance b_s. I assume that the government balances its budget in every period. Each period, total unemployment insurance is equal to

$$\sum_{\mathbf{s}_w} b_s u(\mathbf{s}_w, \mathbf{s}_f, \omega), \tag{12}$$

where $u(\mathbf{s}_w, \mathbf{s}_f, \omega)$ is the mass of type-$\mathbf{s}_w$ unemployed job searchers in submarket $(\mathbf{s}_w, \mathbf{s}_f, \omega)$. Hence, the government's budget constraint is given by:

$$\tau \sum_{\mathbf{s}_w,\mathbf{s}_f,\omega} [1 - u(\mathbf{s}_w, \mathbf{s}_f, \omega)] = \sum_{\mathbf{s}_w} b_s u(\mathbf{s}_w, \mathbf{s}_f, \omega). \tag{13}$$

4.5 Equilibrium

A stationary equilibrium for this economy consists of a set of value functions $\{V_E, V_U\}$ for workers; value functions $\{J_F, \Pi\}$ for firms; search and education policy functions $\{(\mathbf{s}_f, \omega), s\}$ for workers; a distribution of matches $\Omega_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$; a distribution of unemployed workers over ability and education $\Omega_U(\mathbf{s}_w)$; a distribution of non-labor force participants $\Omega_D(a)$; taxes $\{\tau\}$; and tightness $\theta(\mathbf{s}_w, \mathbf{s}_f, \omega)$ for each submarket such that:

1. *Worker Optimization:* The value functions $V_E(\mathbf{s}_w, \mathbf{s}_f, \omega)$ and $V_U(\mathbf{s}_w)$ solve (9), (10), and (11), with associated policy functions $\{s(a), \omega(a, s(a)), x(a, s(a)), z(a, s(a)), k(a, s(a))\}$.

2. *Firm Optimization:* The value function $J_F(\mathbf{s}_w, \mathbf{s}_f, w)$ solves (7).

3. *Free Entry:* Firms enter submarkets until the zero-profit condition (8) holds.

4. *Government Budget Balance:* The government's budget constraint (13) holds with equality in every period.

5. The distributions of workers across employment states $(\Omega_E(\mathbf{s}_w, \mathbf{s}_f, \omega), \Omega_U(\mathbf{s}_w), \Omega_D(a))$ are stationary.

5 Calibration

The model solution consists of an initial and final steady state. To calibrate the steady states, I drawn on some values from the literature, calculate others directly from data, and choose the remainder to match salient features of the Canadian economy in 1980 and the early 2000s.

5.1 Functional Forms

The only functional form required to specify prior to calibration is the matching function. For this, I follow Den Haan, Ramey, and Watson (2000) and use a constant returns to scale matching function of the following form:

$$m(u,v) = \frac{u \cdot v}{(u^{\xi} + v^{\xi})^{1/\xi}}, \tag{14}$$

where ξ is the matching elasticity, u is the mass of unemployed workers, and v is the associated mass of vacancies in a given submarket.

5.2 Externally-Calibrated Parameters

Table 6 below lists the parameters determined outside of the model. I assume that all externally-calibrated parameters remain constant between the two steady states.

A period in the model corresponds to one-quarter in the data. As a result, I set the discount factor β to 0.987, implying an aggregate interest rate of 5 percent annually in steady state. The exogenous rate of job separation, η_s, varies by skill group: for high-skill workers, I use a value of $\eta_H = 0.0238$ and, for low-skill workers, I use a value of $\eta_L = 0.0501$. The labor market exit rate, δ, is set to 0.0063, so that the average working life is 40 years. The matching elasticity is set to 1.60, as in Schaal (2017). The shape parameter for the Extreme Value Type-I distribution governing the taste shocks is set to 0.40 to ensure an interior solution. Finally, as a normalization, I set $h_0 = 1$.

Table 6: Parameters determined outside of the model

	Parameter	**Value**	**Source**
β	Discount factor	0.987	Risk-free rate of 5%
η_{s_0}	Separation rate (Low-skill)	0.050	Flinn and Mullins (2015)
η_{s_1}	Separation rate (High-skill)	0.024	Flinn and Mullins (2015)
ξ	Matching elasticity	1.600	Schaal (2017)
δ	Exit rate	$\frac{1}{160}$	Average working life of 40 years
h_{s_0}	Productivity from base education	1.000	Normalization

5.3 Internally-Calibrated Parameters

The remaining parameters of the model are calibrated in two stages. In the first stage, I calibrate the entire set of parameters to empirical averages from the WES micro-data. This exercise yields a solution to the final steady state of the model. In the second stage, I fix a set of 7 parameters, $\{\kappa, h_{s_1}, \varepsilon_x(s_1), \varepsilon_x(s_0), \varepsilon_k(s_0), c_k(s_0), c_k(s_1))\}$, to their final steady state values and vary the remaining 4 parameters, $\{c_s, c_x(s_0), c_x(s_1), \varepsilon_k(s_1)\}$, such that the share of high-skill workers and training participants by education in the initial steady state match their 1980 empirical averages, and the model-implied increase in the college wage premium is consistent with the actual increase between 1980 and 2000.[14] Table 7 reports the results of the calibration exercise.

As indicated in the third panel of Table 7, I set the vacancy posting cost κ to match an average unemployment rate of 7 percent, which was obtained using publicly-available aggregate information on the unemployment rate from the OECD. I allow unemployment insurance, b_s, to vary across education groups and set b_0 and b_1 such that unemployment insurance covers 40% of the average earnings for low- and high-skill workers in the final steady state, respectively. The productivity gain earned from education, captured by h_{s_1}, is set to match a college wage premium of 1.37. The productivity gain earned from training, captured by $\varepsilon_x(s_1)$ and $\varepsilon_x(s_0)$, are set to match the mean earnings premium accrued by low- and high-skill training participants, respectively, which in the data amount to 1.14 and 1.16. The productivity gain from high-tech employment for low-skill workers, captured by $\varepsilon_k(s_0)$, is set to match the average wage premium of 1.19 for low-skill workers employed by high-tech firms. Finally, the setup costs associated with the high-technology varies by education to ensure that both types of workers are hired by high-technology firms, as in the data. The parameters $c_k(s_0)$ and $c_k(s_1)$ are set to match the average share of low- and high-skill employees at high-technology firms, respectively.

The parameters reported in the top and middle panels of Table 7 vary across steady states. These parameters govern the cost of post-secondary education c_s, the cost of training provision by education $c_x(s)$, and the productivity gain from high-tech employment for high-skill workers—that is, the complementarity between high-skill labor and technology. In both cases, the cost parameters are set to match the shares of high-skill workers and training participants by education in the given years. However, since the WES only includes information on high-technology employment for the final steady state, I am required to follow a different procedure for calibrating the complementarity parameter $\varepsilon_k(s_1)$. For the final steady state, I compute an average wage premium of 1.28 for high-skill workers in high-technology firms directly from the WES micro-data. For the initial steady state, I set $\varepsilon_k(s_1)$ to match an average increase in the college wage premium of 12.5 percent, given a college premium in the final steady state of 1.37. The targeted value of 12.5 percent is chosen as an

[14] The share of high-skill workers in 1980 comes from publicly-accessible OECD data. The share of training participants by education in 1980 comes from Statistics Canada (2001) using the 1983-1985 Adult Education and Training Survey (AETS). Finally, the average increase in the college wage premium is taken to be an intermediate value across existing estimates.

intermediate value between the estimate of 0 percent reported in Kryvtsov and Ueberfeldt (2009) and 25 percent reported in Boudarbat, Lemieux, and Riddell (2010).

Table 7: Parameters determined jointly in equilibrium

Parameter		Value	Target	Data	Model
1980 Calibration:					
Education cost	c_s	90.00	Share of high-skill workers	0.23	0.23
Low-skill training cost	$c_x(s_0)$	10.91	Share of low-skill training participants	0.11	0.13
High-skill training cost	$c_x(s_1)$	3.04	Share of high-skill training participants	0.35	0.36
High-skill prod. from high-tech	$\varepsilon_k(s_1)$	1.06	Average increase in the college wage premium	0.13	0.10
2000 Calibration:					
Education cost	c_s	75.86	Share of high-skill workers	0.68	0.67
Low-skill training cost	$c_x(s_0)$	9.07	Share of low-skill training participants	0.26	0.27
High-skill training cost	$c_x(s_1)$	3.45	Share of high-skill training participants	0.42	0.45
High-skill prod. from high-tech	$\varepsilon_k(s_1)$	1.07	Average high-skill high-tech premium	1.28	1.20
Both Years:					
Vacancy posting cost	κ	3.48	Average unemployment rate	0.07	0.06
Productivity from education	h_{s_1}	1.09	Average college wage premium	1.37	1.36
Low-skill prod. from training	$\varepsilon_x(s_0)$	1.05	Average low-skill training premium	1.16	1.17
High-skill prod. from training	$\varepsilon_x(s_1)$	1.01	Average high-skill training premium	1.14	1.19
Low-skill prod. from high-tech	$\varepsilon_k(s_0)$	1.07	Average low-skill high-tech premium	1.19	1.16
Low-skill cost of high-tech	$c_k(s_0)$	2.05	Share of low-skill high-tech workers	0.36	0.36
High-skill cost of high-tech	$c_k(s_1)$	16.60	Share of high-skill high-tech workers	0.57	0.58

Notes: The targets for all parameters in the middle (2000 Calibration) and bottom (Both Years) panels are drawn from the Workplace and Employee Survey (WES). For the 1980 calibration, the share of high-skill workers is taken from publicly-available OECD data, training participation rates by education are taken from Statistics Canada (2001), and the average increase in the college wage premium is an intermediate value of the estimates reported by Kryvtsov and Ueberfeldt (2009) and Krueger et al. (2010).

6 Quantifying the Importance of Training

The calibrated model provides an exact decomposition of the increase in the college wage premium into three parts: (1) changes in the relatively complementarity between skills and technology, which is governed by the parameter $\varepsilon_k(s_1)$; (2) changes in the share of high-skill workers, which is governed by the parameter c_s; and (3) changes in the shares of training participants by education, which are governed by the parameters $c_x(s_0)$ and $c_x(s_1)$. Accordingly, in this section, I use the model to perform a decomposition analysis and, ultimately, quantify the contribution of training on the rise in the college wage premium.

To this end, I start in the initial steady state and sequentially allow each subset of parameters to adjust to their final steady state values. Following each simulation, I compute the model-implied level of, and change in, the college wage premium and measure their difference from the counterparts computed in the previous step. Finally, I compute the relative contribution of each channel by

Table 8: Decomposition Results

	Setting			
	(1)	**(2)**	**(3)**	**(4)**
Technology-Skill Complementarity, ε_k	–	✓	✓	✓
Education Cost, c_s	–	–	✓	✓
Training Costs, c_x	–	–	–	✓
College wage premium (%)	27.78	31.51	32.22	30.50
College wage premium (Ratio)	1.32	1.37	1.38	1.36
Share of high-skill workers	0.23	0.32	0.71	0.67
Share of high-skill trained workers	0.36	0.95	0.53	0.45
Share of low-skill trained workers	0.13	0.04	0.00	0.27
Share of high-skill high-tech workers	0.58	0.97	0.56	0.58
Share of low-skill high-tech workers	0.36	0.65	0.28	0.36

– indicates that the parameter value is held fixed at initial (1980) value.
✓indicates that the parameter value varies across steady states.

comparing the estimated difference at each step to the overall change in the premium.[15] The results from performing this exercise are summarized in Table 8.

Consider first columns (1) and (2) of Table 8. The difference in moments reported between the two columns solely reflects changes in the relative complementarity between high-skill labor and technology, that is, $\varepsilon_k(s_1)$. Relative to the initial steady state, the college wage premium in this environment increases by 3.73 percentage points, or 13.4 percent over the 1980 level. The *increase* in the college wage premium therefore overshoots the model-attained value of 9.79 percent by an additional 37 percent. This phenomenon occurs because the increase in technology-skill complementarity encourages firms to invest in the high-technology and raise their demand for high-skill workers. Combined with the positive association between technology and training, high-skill workers experience an extra boost in the probability of receiving training on the job. Therefore, increased labor demand and training participation for high-skill workers jointly act to raise the college wage premium.

The difference in values reported in columns (2) and (3) of Table 8 reflect the joint impact of changes in technology-skill complementarity and the costs of post-secondary education. In this environment, the college wage premium increases even further from 31.51 percent to 32.22 percent, implying an increase of 16 percent over the 1980 level. Although the lower cost of education generates a much larger equilibrium share of high-skill workers, its effect on the college wage premium is completely undone by the response of training. In particular, the large increase in the supply of high-skill workers encourages new entrant firms to shift from posting vacancies with training to posting vacancies for high-skill workers. In the new equilibrium, the training gap widens as virtually

[15]Because training has a dampening effect on the premium, I take the sum of the absolute changes as the measure of the overall change.

no low-skill workers participate.

Finally, Column (4) of Table 8 reports the estimated moments from the final steady state—that is, after allowing all 4 parameters to vary across steady states. By comparing the difference in moments between columns (3) and (4), I obtain an measure of how much changes in training costs have dampened the increase in the college wage premium. In particular, with training costs held fixed (Column (3)), the college wage premium increases by 16 percent over its 1980 value. Once training costs are allowed to vary (Column (4)), the college wage premium increases by only 9.79 percent over its 1980 value. Therefore, absent changes in training costs, the increase in the college wage premium would have been 63 percent larger between 1980 and the early 2000s. On the other hand, the measured *level* of the college wage premium would have been 6 percent larger in the final steady state.

Across all experiments, the level of the college wage premium experiences an absolute change of 8.33 percentage points. Most of this change—approximately 60 percent—is driven by changes in technology-skill complementarity, indicating that technological change is indeed the driving force of the college wage premium. Changes in training account for the next largest share—28 percent—and, in fact, dampen the premium. This finding is consistent with recent work by Doepke and Gaetani (2020), who find increased on-the-job human capital accumulation among low-skill workers helps to explain the relatively subdued growth in the college wage premium experienced in Germany compared to the United States.

7 Conclusion

A large quantitative literature has established a tight link between the introduction of new technologies and rising college wage premia around the globe. In this paper, I use matched employer-employee survey data from Canada to highlight that the diffusion of technological change across firms has also generated a meaningful increase in training participation among workers—particularly, low-skill workers—which, because of the earnings premium associated with training, has dampened the college wage premium over time.

To reach this conclusion, I start by documenting a new set of facts using the Canadian microdata. I show that (1) high-technology firms tend to be relatively more productive, provide more training, and hire more high-skill employees than low-technology firms; (2) training participants earn between 3 to 14 percent more on hourly earnings relative to non-participants; and (3) relative increases in training participation among low-skill workers reduce the college wage premium.

I use these features of the Canadian data to discipline a quantitative model of the labor market and generate the structure necessary to analyze the main drivers of the college wage premium. Consistent with existing evidence, I find that technology-skill complementarity accounts for the greatest share (60 percent) of absolute variation in the college premium over time. Importantly, however, I also find that a relative expansion in training participation among low-skill workers played a crucial role in dampening the college wage premium between 1980 and the early 2000s. In particular, absent changes in training costs, the rise in the college wage premium would have been 63 percent larger over this period. This latter finding suggests that potential gains in equity between high- and low-skill workers may be recovered if policymakers develop programs to encourage training participation among low-skill workers.

References

Açıkgöz, Ömer Tuğrul and Barış Kaymak (2014). "The Rising Skill Premium and Deunionization". *Journal of Monetary Economics*, 63, 37–50.

Amiti, Mary and Lisa Cameron (2012). "Trade Liberalization and the Wage Skill Premium: Evidence from Indonesia". *Journal of International Economics*, 87 (2), 277–287.

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.

Bartel, Ann P. and Nachum Sicherman (1998). "Technological Change and the Skill Acquisition of Young Workers". *Journal of Labor Economics*, 16 (4), 718–755.

Black, Sandra E. and Lisa M. Lynch (2004). "What's Driving the New Economy?: The Benefits of Workplace Innovation". *The Economic Journal*, 114 (493), F97–F116.

Blundell, Richard, Lorraine Dearden, and Costas Meghir (1996). *The Determinants and Effects of Work-Related Training in Britain*. R50. IFS report.

Blundell, Richard, David A. Green, and Wenchao Jin (2022). "The U.K. as a Technological Follower: Higher Education Expansion and the College Wage Premium". *Review of Economic Studies*, 89 (1), 142–180.

Boothby, Daniel, Anik Dufour, and Jianmin Tang (2010). "Technology Adoption, Training and Productivity Performance". *Research Policy*, 39 (5), 650–661.

Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell (2010). "The Evolution of the Returns to Human Capital in Canada, 1980–2005". *Canadian Public Policy*, 36 (1), 63–89.

Bound, John and George Johnson (1992). "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations". *American Economic Review*, 82 (3), 371–392.

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.

Burbidge, John B., Lonnie Magee, and A. Leslie Robb (2002). "The Education Premium in Canada and the United States". *Canadian Public Policy*, 203–217.

Burstein, Ariel and Jonathan Vogel (2017). "International Trade, Technology, and the Skill Premium". *Journal of Political Economy*, 125 (5), 1356–1412.

Card, David and John E. DiNardo (2002). "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles". *Journal of Labor Economics*, 20 (4), 733–783.

Den Haan, Wouter J., Garey Ramey, and Joel Watson (2000). “Job Destruction and Propagation of Shocks”. *American Economic Review*, 90 (3), 482–498.

Dix-Carneiro, Rafael and Brian K. Kovak (2015). “Trade Liberalization and the Skill Premium: A Local Labor Markets Approach”. *American Economic Review*, 105 (5), 551–557.

Doepke, Matthias and Ruben Gaetani (2020). “Why Didn’t the College Premium Rise Everywhere? Employment Protection and On-the-Job Investment in Skills”. *NBER Working Paper Series*, (27331).

Flinn, Christopher, Ahu Gemici, and Steven Laufer (2017). “Search, Matching, and Training”. *Review of Economic Dynamics*, 25, 260–297.

Flinn, Christopher and Joseph Mullins (2015). “Labor Market Search and Schooling Investment”. *International Economic Review*, 56 (2), 359–398.

Glitz, Albrecht and Daniel Wissmann (2021). “Skill Premiums and the Supply of Young Workers in Germany”. *Labour Economics*, 72, 102034.

He, Hui (2012). “What Drives the Skill Premium: Technological Change or Demographic Variation?” *European Economic Review*, 56 (8), 1546–1572.

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.

Kristal, Tali (2020). “Why Has Computerization Increased Wage Inequality? Information, Cccupational Structural Power, and Wage Inequality”. *Work and Occupations*, 47 (4), 466–503.

Krueger, Dirk et al. (2010). “Cross-sectional facts for macroeconomists”. *Review of Economic dynamics*, 13 (1), 1–14.

Krusell, Per et al. (2000). “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis”. *Econometrica*, 68 (5), 1029–1053.

Kryvtsov, Oleksiy and Alexander Ueberfeldt (2009). *What Accounts for the US-Canada Education-Premium Difference?* Tech. rep. Bank of Canada.

Leuven, Edwin and Hessel Oosterbeek (2008). “An Alternative Approach to Estimate the Wage Returns to Private-Sector Training”. *Journal of Applied Econometrics*, 23 (4), 423–434.

Lindner, Attila et al. (2022). “Firm-Level Technological Change and Skill Demand”. *Centre for Economic Performance Working Paper*, (1857).

Lise, Jeremy et al. (2014). “Wage, Income and Consumption Inequality in Japan, 1981–2008: From Boom to Lost decades”. *Review of Economic Dynamics*, 17 (4), 582–612.

Matsuda, Kazushige (2020). “Optimal Timing of College Subsidies: Enrollment, Graduation, and the Skill Premium”. *European Economic Review*, 129, 103549.

Menzio, Guido and Shouyong Shi (2010). "Block Recursive Equilibria for Stochastic Models of Search On the Job". *Journal of Economic Theory*, 145 (4), 1453–1494.

— (2011). "Efficient Search On the Job and the Business Cycle". *Journal of Political Economy*, 119 (3), 468–510.

Mincer, Jacob (1958). "Investment in Human Capital and Personal Income Distribution". *Journal of Political Economy*, 66 (4), 281–302.

Moen, Espen R. and Åsa Rosén (2004). "Does Poaching Distort Training?" *The Review of Economic Studies*, 71 (4), 1143–1162.

Parent, Daniel (1999). "Wages and Mobility: The Impact of Employer-Provided Training". *Journal of Labor Economics*, 17 (2), 298–317.

Parro, Fernando (2013). "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade". *American Economic Journal: Macroeconomics*, 5 (2), 72–117.

Salgado, Sergio (2020). "Technical Change and Entrepreneurship". *SSRN 3616568*.

Schaal, Edouard (2017). "Uncertainty and unemployment". *Econometrica*, 85 (6), 1675–1721.

Shephard, Andrew and Modibo Sidibe (2019). "Schooling Investment, Mismatch, and Wage Inequality". *PIER Working Paper*.

Shi, Shouyong (2002). "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology". *Review of Economic Studies*, 69 (2), 467–491.

Statistics Canada (2001). *A Report on Adult Education and Training in Canada: Learning A Living*. Government of Canada, Ottawa, Ontario.

Takahashi, Shuhei and Ken Yamada (2022). "Understanding International Differences in the Skill Premium: The Role of Capital Taxes and Transfers". *Journal of Economic Dynamics and Control*, 143, 104511.

Velden, Rolf Van Der and Ineke Bijlsma (2016). "College Wage Premiums and Skills: A Cross-Country Analysis". *Oxford Review of Economic Policy*, 32 (4), 497–513.

Walker, Ian and Yu Zhu (2008). "The College Wage Premium and the Expansion of Higher Education in the U.K." *The Scandinavian Journal of Economics*, 110 (4), 695–709.

Zentler-Munro, David (2021). "Rising Wage Inequality: Technological Change and Search Frictions". *Working Paper*, 1–60.

A Data Sources

A.1 Workplace and Employee Survey (WES)

The main data source for this paper is the Workplace and Employee Survey (WES). The WES is a matched employer-employee survey data set from Canada, which covers approximately 20,000 employees and 6,000 employers at an annual frequency from 1999 to 2006. In each year, the WES contains two components: a workplace component and an employee component. The target population for the workplace component consists of all business locations operating in Canada with paid employees in March of the survey year with the exception of employers operating in the Territories; crop or animal production; fishing, hunting, and trapping; private households, religious organizations, or public administration. Hence, each observation in the workplace component of the WES is an *establishment*. In the main text, I use the words "workplace", "establishment", "employer", and "firm" interchangably. The target population for the employee component of the WES consists of all employees working or on paid leave in March of the survey year who (1) are employed by an establishment in the workplace component and (2) receive a Canada Revenue Agency T-4 Supplementary form. Workers receiving T-4 slips from multiple different workplaces are counted as distinct observations in the employee component of the WES.

The sampling methodology of the WES is divided into two parts. First, a sample of employers is drawn from the Business Register at Statistics Canada. Second, employees from the participating workplaces are selected at random from lists provided by employers to the surveyors. The initial sample was drawn in 1999. Every two years, a subset of new employers is added to the workplace component from establishments added to the Business Register since the last survey occasion. Every two year, a new sample of employees from the participating workplace is also drawn. Hence, the employee component of the WES is fully refreshed every odd year, while the workplace component is only partially refreshed. I pool the cross-sectional data from the 1999, 2001, 2003, and 2005 surveys, and restrict attention to workers of age 25 to 64 years old.

The main variables of interest from the WES include: educational attainment, classroom and on-the-job training participation, and earnings on the employee side; and, firm size, the share of employees using computers, and classroom or on-the-job training provision on the workplace side. Details about how I construct the measures of skills, training, and technology used in the empirical analysis are contained in the main text.

A.2 Additional Data Sources for Calibration

For the calibration of the initial steady state, I use two additional data sources. The first is publicly-available aggregate data from the OECD, which is available here. I use the OECD data to obtain a target for the average unemployment rate and share of high-skill (tertiary-educated) workers in 1980. The second source is Statistics Canada (2001), which documents training participation rates by education from the past revisions of the Adult Education and Training Survey (AETS).

B Empirical Specifications

The regression specifications underlying Fact 1 of Section 3 in the main text are detailed below. In all cases, I estimate the standard errors by bootstrap using the bootstrap weights provided by Statistics Canada and 100 replications.

B.1 Workplace-Level Productivity

In the first regression, I estimate the effect of technology on (log) revenue productivity according to the following specification:

$$\ln(\text{Productivity}_{j,t}) = \beta_0 + \beta_1 \text{HighTech}_{j,t} + \xi Z_{j,t} + \theta_t + \varepsilon_{j,t}. \tag{15}$$

where j indexes firms and t indexes time. The main covariate of interest is $\text{HighTech}_{j,t}$, which is an indicator equal to 1 if firm j is a high-tech firm and 0 if low-tech. I also include a vector $Z_{j,t}$ of time-varying workplace-level control variables, which includes industry, training provision, and firm size; a set of year-fixed effects θ_t; and an error term $\varepsilon_{j,t}$.

B.2 Employee-Level Training Participation

In the second regression, I estimate the probability of training participation at the employee-level to evaluate whether employees of high-technology firms are relatively more or less likely to receive training. To this end, I estimate a logistic regression of the following form:

$$\Pr[\text{Train}_{i,t} = 1] = \beta_0 + \beta_1 \text{HighTech}_{j,t} + \beta_2 \text{HighSkill}_{i,t} + \delta X_{i,t} + \xi Z_{j,t} + \theta_t + \varepsilon_{i,t}, \tag{16}$$

where, again, i indexes individuals, j indexes firms, and t indexes time. The main covariate of interest is the indicator $\text{HighTech}_{j,t}$, which equals 1 if worker i's employer j is a high-technology firm. I also control for the worker's level of education, $\text{HighSkill}_{i,t}$, a set of time-varying employee-level covariates $X_{i,t}$, a set of time-varying workplace-level covariates $Z_{j,t}$, year fixed-effects θ_t, and an error term $\varepsilon_{i,t}$.

B.3 Probability of High-Technology Employment

For the third regression, I estimate the probability of being employed by a high-technology firm. That is, for each individual i employed by workplace j in year t, I estimate the following logistic regression:

$$\Pr[\text{HighTech}_{j,t} = 1] = \beta_0 + \beta_1 \text{HighSkill}_{i,t} + \delta X_{i,t} + \xi Z_{j,t} + \theta_t + \varepsilon_{i,t}, \tag{17}$$

where $\text{HighTech}_{j,t}$ indicates whether employee i's workplace j is high-technology; $\text{HighSkill}_{i,t}$ is an indicator for whether employee i is a high-skill worker; $X_{i,t}$ is a vector of time-varying worker-level control variables, which includes training participation, occupation, age, experience, gender, CBA coverage, immigration status, and tenure; and $Z_{j,t}$ is a vector of time-varying workplace-level control variables, which includes industry, firm size, and productivity.