Canadian Labour Market and Skills Researcher Network

Working Paper No. 45

The Impact of Aggregate and Sectoral Fluctuations on Training Decisions

Vincenzo Caponi
Ryerson University

Cevat Burc Kayahan
Acadia University

Miana Plesca
University of Guelph

October 2009

CLSRN is supported by Human Resources and Social Development Canada (HRSDC) and the Social Sciences and Humanities Research Council of Canada (SSHRC).

The Impact of Aggregate and Sectoral Fluctuations on Training Decisions*

Vincenzo Caponi[†], Cevat Burc Kayahan[††] and Miana Plesca[§]

Abstract

The literature has not yet resolved whether the effect of macroeconomic fluctuations on training decisions is positive or negative. On the one hand, the opportunity cost to train is lower during downturns, and thus training should be counter-cyclical. On the other hand, a positive shock may be related to the adoption of new technologies and increased returns to skill, making training incidence pro-cyclical. Using the Canadian panel of Workplace and Employee Survey (WES), we document another important channel at work: the relative position of a sector also matters. We find not only that training moves counter-cyclically with the aggregate business cycle (more training during downturns), but also that the idiosyncratic sectoral shocks have a positive impact on training incidence (more training in sectors doing relatively better). These findings help us better understand training decisions by firms.

JEL Classification: J24, E32
Keywords: Training; Human capital; Business cycles; Sectoral shocks.

---

*This research was financially supported by a grant from the Canadian Labour Research Network (CLSRN). We thank Louise Grogan, Gueorgui Kambourov, Chris Robinson, seminar participants at the University of Guelphand Mount Allison University, as well as participants at the CLSRN 2008 Workshop, the 2008 Atlantic Canada Economic meetings, and the 2009 Canadian Economic Association meetings, for very helpful discussions.

[†]Department of Economics, Ryerson University, Toronto, Canada. IZA, Institute for the Study of Labour, Bonn, Germany. RCEA, Rimini Center for Economic Analysis, Rimini, Italy. Email: vcaponi@ryerson.ca.

[††]Department of Economics, Acadia University, Wolfville, Nova Scotia, Canada. Email: ckayahan@acadiau.ca

[§]Department of Economics, University of Guelph, Guelph, Ontario, Canada. Email: miplesca@uoguelph.ca.

Executive Summary

Does training rise or fall with economic fluctuations? On the one hand, a negative productivity shock may be associated with increased training, since the opportunity cost to train workers is lower in downturns. On the other hand, a positive shock may be related to the adoption of new technologies which may require training and can provide increased returns to skill. Both the counter-cyclical and pro-cyclical arguments have sound theoretical justifications, yet we have little evidence which one will hold true or dominate the other empirically.

Our contribution is to provide a unifying framework where two channels coexist. We bring empirical evidence that training is counter-cyclical - as expected, the aggregate output shock has a negative impact on the incidence of firm training. More importantly, we show that the idiosyncratic sectoral shocks are pro-cyclical - firms from sectors which experience a positive shock relative to the rest of the economy have an incentive to train more, even after conditioning for innovation and adoption of new technologies. We contend, and bring some related evidence, that sectors doing relatively better attract workers from sectors doing relatively worse, and these workers may require remedial training in specific skills.

To measure the effect of aggregate and sectoral output fluctuations on training incidence we use the panel of the Canadian Workplace and Employee Survey (WES) together with statistics on industrial output. We consider two main definitions for training by firms: a binary indicator whether the firm has provided training or not, which we call "extensive margin of training," and a continuous measure of training which we call "intensive margin" (conditional on firms who train) expressed either as a percentage of workforce trained or as training expenditures per worker. Our major findings are that (i) training moves counter-cyclically with the aggregate output fluctuations (more training in downturns), while at the same time (ii) the relative position of sectoral output has a positive impact on training decisions (more training in a sector doing relatively better). The magnitude of these two channels is comparable. We find that a one-percentage point increase in the deviation of aggregate output relative to its trend decreases the probability of training by 1.5 percentage points and decreases training expenditures by $7 per worker. A one-percentage point increase in the share of a sector's output – controlling for aggregate shocks and for innovation and adoption of new technologies − increases the probability of training by 0.7 percentage points and increases training expenditures per worker by $19 for the firms who train.

The finding that two opposing channels through which output fluctuations affect training decisions is relevant for at three main reasons: (i) it gives us better insights into firms' training decisions over the business cycle, (ii) it quantifies how aggregate and sectoral shocks play into the human capital accumulation decision, and (iii) it helps policy-makers understand that fluctuations in training incidence that may be an optimal response to macroeconomic shocks, and not necessarily an indicator of underinvestment in training.

From a theoretical standpoint, this study highlights the importance for any models of firm training to incorporate mechanisms coming from both aggregate and sectoral output fluctuations. To frame our empirical results, we add training to the Mortensen-Pissarides search model. The equilibrium adjustments with aggregate and sectoral shocks illustrate the mechanisms that are highlighted in our empirical analysis. The aggregate and sectoral output fluctuations documented here can help inform policy-makers on whether observed trends in training are dictated by economic circumstances, or whether firms under-invest in training, which may suggest an area where government policy intervention can be explored.

1 Introduction

Human capital increases through training, be that implicit on-the-job training (as captured by tenure), or explicit classroom-type formal training leading to human capital accumulation. Our focus in this paper is mostly on the latter: we are interested in how a firm's explicit decision to train depends on aggregate and sectoral output fluctuations.

It is not ex-ante obvious whether investments in human capital are counter-cyclical, pro-cyclical, or a-cyclical. Since the opportunity cost to train workers is lower in downturns, a negative productivity shock should be associated with increased training. This channel is highlighted by deJong and Ingram (2001) who find that training activities "are distinctively countercyclical" and by Devereux (2000) who argues that during downturns firms hoard labour by assigning high-skill workers to lower-production activities such as training, thus avoiding layoffs and the fixed costs associated with firing and re-hiring workers. On the worker side, the literature documents that college enrollment is counter-cyclical (e.g. Dellas and Sakellaris (2003)); typically, enrollment in universities increases when the economy is not doing well and good jobs are harder to find.

Nevertheless, a different adjustment is also possible. A positive shock may be related to the adoption of new technologies which not only require training but also can provide increased returns to skill. This is the channel identified by King and Sweetman (2002). Using administrative Canadian data, they find that "re-tooling" (measured as quits from work to school) is pro-cyclical, consistent with a model where the outside option of high-skill jobs goes up during episodes of higher output, increasing the return to skill and therefore the value of training.

Arguments for both counter-cyclical and pro-cyclical investments in skill have sound theoretical justifications, yet we have little evidence which one should dominate empirically. Our contribution here is to provide a framework where two channels, potentially opposed, coexist. On the one hand, we find training to be counter-cyclical: the aggregate output shock has a negative impact on the incidence of firm training. Like most of the literature, we contend this channel is working through the opportunity cost of training (foregone output) which is lower in downturns. On the other hand, we identify a new channel that relates training with idiosyncratic sectoral shocks. Firms from sectors experiencing a positive shock relative to the rest of the economy have an incentive to train more, while firms from sectors experiencing a relatively negative shock have an incentive to train less. We contend there are two mechanisms at work here. First, insofar as positive sectoral shocks are related to the adoption of new technologies, firms will invest in training to operate these new technologies. Second, workers reallocate from the sectors doing relatively worse into the ones doing relatively better; the workers who are new to a sector may require training in sector-specific skills. Indeed, we document empirically that this is the case: (i) firms who innovate or adopt new technologies train more; (ii) there is an increase in training incidence by firms in sectors doing relatively better, even after controlling for the adoption of new technologies, and (iii) the probability of a worker to get trained is higher when the worker is new to a sector.

To measure the effect of aggregate and sectoral output fluctuations on training incidence, we use the Canadian Workplace and Employee Survey (WES). The WES is an eight-year panel of firms (1999 to 2006) representative of all industries except agriculture.[2] It is particularly appealing because response rates are consistently high across all panel years, and sample sizes are relatively generous. Most importantly, the panel nature of the WES allows us to remove the unobserved firm-specific fixed effects in the empirical analysis. Finally, the WES is a linked employer-employee survey where, as long as she stays with the original firm, the worker can be surveyed for two consecutive years. While our analysis focuses mostly on the firm, we also analyze information from the worker side as supporting evidence for sectoral reallocation and worker training. There is one caveat: since very few workers are interviewed per establishment, they are not representative of a firm's workforce; nevertheless, they are representative of the overall economy.

We consider two main definitions for training by firms: a binary indicator whether the firm has provided training or not, which we call "extensive margin of training," and a continuous measure of training which we call "intensive margin" (conditional on firms who train) expressed either as a percentage of workforce trained or as training expenditures per worker. Our major findings are that (i) training moves counter-cyclically with the aggregate output fluctuations (more training in downturns), while at the same time (ii) the relative position of sectoral output has a positive impact on training decisions (more training in a sector doing relatively better).

[2]The unit of analysis is the establishment. Throughout the paper, we refer to it interchangeably as "firm" or "workplace."

The magnitude of these two channels is comparable. We find that a one-percentage point increase in the deviation of aggregate output relative to its trend *decreases* the probability of training by 1.5 percentage points and decreases training expenditures by \$7 per worker. A one-percentage point increase in the share of a sector's output − controlling for aggregate shocks − *increases* the probability of training by 0.7 percentage points and increases training expenditures per worker by \$19 for the firms who train. These results are robust to different specifications.

The paper proceeds as follows. Section 2 describes the microdata used in the analysis and references the sources of data for sectoral and aggregate output. The main results are in Section 3 where we describe the empirical estimation methodology and our findings. Section 4 presents a theoretical model useful to interpret our results. Theory and policy implications are discussed in Section 5, which concludes.

2 Data

2.1 WES Data

We use the Canadian Workplace and Employee Survey (WES) from 1999 to 2006, a nationally representative matched employer-employee survey with a longitudinal design. The WES targets all workplaces in Canada with paid employees in March.[3] The sample of locations in the frame is stratified by industry (14 strata), region (6 strata), and firm size (3 strata). The stratification of units remains constant over the life of the initial panel; survey weights are used throughout the analysis. The response rate for the workplace side is 74.9% for the 1999-2006 period, which is relatively high for a panel survey of firms.

The linked employee component of WES is based on lists of employees made available by the selected workplaces. A maximum of twenty four employees are sampled. In workplaces with fewer than four employees, all employees are selected. Employees are surveyed for at most two consecutive years, after which they are dropped from the sample and replaced with other employees. While workers can be linked to their firms, it can be difficult to infer firm-specific distributions from the worker side, since only a few workers (sometimes as little as three) are

[3] There are some exceptions. Certain industries are not sampled: farming, public administration, and religious organizations. Also excluded are remote locations: Yukon, Nunavut, and the Northwest Territories.

interviewed per establishment. Nevertheless, the worker side of WES is important for our analysis by allowing us to investigate how the tenure of a worker − within the firm and, more importantly, within the sector − influences the propensity for the worker to get trained.

We define the "extensive margin" of training (based on firms' self-reported answers) as an indicator whether the firm had offered any training to its workers. While we focus our attention on formal classroom training (CT) as a tool for human capital accumulation, we also present sensitivity results to using a broader measure of skill improvement which combines classroom training with on the job training (CT+OJT).[4] The intensive margin of training comes either from the percentage of workforce offered formal classroom training (CT) by each firm, or from the firm's training expenses per worker; we report analyses on both. The means of the training variables are presented in the top panel of Table 1. To control for observed firm-specific determinants of training, we follow the literature (*e.g.*, Turcotte, Leonard, and Montmarquette (2003)) by using firm size, innovation, unionization, output market and workforce skill distribution. These variables are listed in the bottom panel of Table 1.

2.2 Output fluctuation series

To capture the aggregate business cycle effects we use the Gross Domestic Product (GDP) series from Statistics Canada. Series for the overall economy, as well as by sectors, are available since the early 1980s and are reported in 2000 constant dollars. Since the time period surveyed by the WES is between April 1^{st} of the previous year and March 31^{st} of the current year, we use quarterly GDP aggregated into annual series to correspond to the timing in the WES. We detrend the real log GDP series using the Hodrick-Prescott filter with a smoothing parameter of $\Lambda = 6.15$. While we experimented with different smoothing parameters all the way up to $\Lambda = 10$, changing the bandwidth did not affect our results. Figure 1 presents the HP-filtered log GDP series.

The classification of sectors in the WES follows for the most part the two-digit North American Industry Classification System (NAICS) with a few small differences: in the WES, some industries from the NAICS are aggregated into a single group, and firms from the agricultural

[4] The WES questionnaire defines classroom training as "all training activities i) which have a pre-determined format including a pre-defined objective ii) which have a specific content, iii) for which progress may be monitored and/or evaluated."

sector are not sampled in the WES. We aggregate the sectors from the output fluctuation statistics in a manner consistent with the WES. The sectors used in the analysis together with their relative shares are listed in Table 2. They are Forestry and Mining, Construction, Transportation, Information and Communication, Finance, Real Estate, Business Services, Education and Health, Manufacturing, and Retail. For the relative position of each sector we use the share of that sector's output in total output.[5]

3 Empirical evidence

The extensive and intensive margin analyses are both relevant; the former refers to a model where firms decide whether to train or not, while the latter refers to a model where selected firms decide how much to train. When the dependent variable is discrete, we estimate a fixed-effects logit model; when the dependent variable is continuous, we estimate a linear fixed-effects model. In all cases, we allow for unobserved firm heterogeneity, which can be differenced out:[6]

$$y_{it} = \alpha_i + \beta X_{it} + u_{it} \qquad t = 1, \dots, T \tag{1}$$

where y_{it} is the training decision of firm i in period t, α_i denotes firm-specific heterogeneity, X_{it} are firm-specific and aggregate characteristics affecting training, and u_{it} represents idiosyncratic errors. The first differences version removes the fixed effect and is very close to the fixed-effects estimator.

$$\Delta y_{it} = \beta \; \Delta X_{it} + \Delta u_{it}. \tag{2}$$

Under the strict exogeneity identification condition:

$$E(u_{it} \,| X_{i1}, X_{i2}, \dots, X_{iT}, \alpha_i = 0, \,) \qquad t = 1, ..., T, \tag{3}$$

the coefficient vector β can be estimated consistently with fixed effect estimators allowing for any correlation between X_{it} and the fixed effect α_i. The strict exogeneity assumption rules out correlations between X_{it+k} and u_{it} (no feedback effect). While this may seem like a strong assumption, nevertheless, given that we explicitly control for shocks at the aggregate and sectoral levels in the vector X, strict exogeneity can be plausible in our case. Moreover, while

[5]For reasons of space, we omit from here graphs with the relative sectoral position and training incidence by respective sectors.

[6]Hausman tests reject random effects in favour of fixed-effects in each specification.

we start by presenting results computed under the strict exogeneity assumption, in Section 3.3 we report results from a more general specification where we instrument for the possible correlation between X_{it+1} and u_{it}.

3.1 The impact of output fluctuations on training: extensive margin

We start by investigating the extensive margin: how do macroeconomic factors influence the binary decision of firms to train or not to train? The results are in Table 3. The first specification defines training as an indicator of whether the firm has provided formal classroom training or not (column 1), and the second specification adds on-the-job training (column 3). The macroeconomic factors are the HP-detrended log GDP and the share of each sector in total output. All results reported here also include the HP-filtered trend as a regressor. Sensitivity analysis (available from the authors) indicates that including the trend or not does not change the results.

While in the linear models the coefficients represent marginal effects, this is no longer the case with the fixed-effects logit model. Nevertheless, the ratio of logit coefficients equals the ratio of marginal effects; in this sense, the quantitative results from the logit models can be directly compared with those from linear models. Moreover, we report in Appendix II similar results from a linear model estimation (OLS instead of logit) to provide some benchmark for the marginal effects corresponding to our coefficients.[7]

The coefficient on aggregate output fluctuations (-0.072) is negative and significant, implying that training is counter-cyclical. In other words, firms are more likely to train their workforce during periods of aggregate output slow-downs. This is in line with the argument that workers are relatively less productive during downturns, and thus the opportunity cost of training (foregone output) is relatively smaller during recessions. By contrast, the coefficient on the relative position of the sector (0.034) is positive and significant: firms in sectors doing relatively better are more likely to train. The magnitude of the sectoral adjustment is roughly half that of the aggregate macro channel.[8]

[7] For non-linear fixed-effects models (logit in our case), the marginal effects are difficult to compute. For each observation, the marginal effect would be $\Lambda[\beta X_{it}\alpha_i](1-\Lambda[\beta X_{it}\alpha_i])$ where Λ is the logistic function. Without explicit distributional assumptions for the unobserved heterogeneity term α_i we cannot average it out across the population (Wooldridge (2005)).

[8] In the linear model estimation − reported in Appendix II,− a percentage-point increase in the ratio of

The impact of the innovation/adoption of new technology indicator is large and positive (0.629), which is as expected. Notably though, the aggregate macro channel and the relative sectoral channel have a significant impact on training decisions, even after controlling for the innovation variable.[9]

A very similar story emerges when the definition of training is expanded to add on-the-job training to the formal classroom training definition from above. The results are in the last two columns of Table 3. Although the magnitude of the two coefficients of interest - aggregate and sectoral fluctuations - is slightly smaller, their signs and relative sizes indicate that training moves counter-cyclically with the business cycle and pro-cyclically with the idiosyncratic sectoral shock. This second sectoral channel, previously ignored in the literature, is of substantive relevance.

3.2 The impact of output fluctuations on training: intensive margin

We move now to investigate the impact of macroeconomic fluctuations on the training decision by considering more detailed training variables: (i) the percentage of workforce trained by a firm, and (ii) the training expenditures per worker.

One of the methodological advantages of a continuous left-hand side variable is that we estimate a linear panel model where coefficients directly translate into marginal effects. Moreover, in the linear model we can account for the potential correlation in the error terms across firms within a sector. Insofar as the i.i.d. assumption is violated, clustering the standard errors across sectors provides for consistent inference. We report standard errors both non-clustered and clustered by sector. Since the robust standard errors are larger, statistical significance

GDP fluctuations to the HP trend will decrease the probability that a firm trains by roughly one and a half percent. A one percentage-point increase in the share of the sector's output in total output will increase the firm's propensity to train by 0.7 percent. The OLS numbers are consistent with the rule of thumb that marginal effects are roughly one quarter to twenty-percent of the logit coefficients. Moreover, the ratio of coefficients gets preserved; for instance, the ratio of aggregate relative to sectoral impacts is roughly 2.1 in both linear and non-linear models.

[9]In terms of the other firm-specific factors that influence training, we find that on average more diversified and larger firms are more likely to train. Firms who are unionized are also more likely to train, except when adding OJT to the CT definition, when the union coefficient becomes negative (note though that the union incidence in our sample is very low, due to the fact that the highly-unionized public sector is excluded from the WES.) Relative to Professionals, increasing employment in any category other than Technical results in more training. This is in line with what has been documented elsewhere in the literature of firm training determinants − see for instance Lynch and Black (1998) for the U.S.; Dearden, Reed, and Reenen (2006) for the U.K.; and Turcotte, Leonard, and Montmarquette (2003) for Canada.

becomes problematic for a lot of the control variables. Nevertheless, the relevant variables for our analysis, aggregate and sectoral fluctuations and innovation, remain significant throughout the analysis.[10]

We start with a specification that includes all firms in the sample, those who train a positive amount as well as those that do not train (results in Table 4). We use this specification to benchmark the results from the intensive margin analysis where we select in the sample only firms who train positive amounts (results in Table 5). While the magnitude of coefficients is obviously larger in the positive training specification, the differences are small and do not change the substance of the analysis.[11]

The first column in Table 4 is for the left-hand side variable defined as the percentage of the workforce trained by the firm. The next two columns report standard errors, without and with accounting for clustering by industries. The second specification reports the results when training is measured as training expenditures per worker (column 4) followed by unclustered and clustered standard errors. A percentage-point increase in the deviation of (detrended) output from HP trend decreases the percentage of workforce trained by .65 percent. A percentage-point increase in the share of a sector in total output increases the percentage of workforce trained by .53 percent. When looking at the impact on training expenditures per worker (column 4), the percentage change in the business cycle measure, GDP fluctuations, decreases the training expenditure by $4, but this impact is not statistically significant. By contrast, an increase in the relative share of a sector by one percentage-point will increase the training expenditures by $14 (statistically significant).

The same picture emerges when we focus on firms that consistently train, only the magnitude of the impacts is larger. These results are in Table 5. Conditional on positive training ("intensive margin"), a percentage-point increase in the GDP deviation decreases the percentage of workforce trained by .895 percent and the training expenditures per worker by $7. A percentage-point increase in the relative position of the sector increases the share of the workforce trained by 0.7 percentage-points and the training expenditure by $19 per worker. Firms

[10]While it may be interesting to revisit the literature of firm-specific determinants of training given how standard errors increase when we cluster by sector, this is not our goal. For the purpose of this paper, we are satisfied that the statistical significance of the main variables, GDP and sectoral deviations, remains unaffected.

[11]Moreover, we obtain largely similar results if we restrict the analysis to firms who train every period.

who innovate and adopt new technologies train 8 percent more of their workforce and spend $76 more per worker compared to firms who do not innovate. All coefficients except for the impact of GDP fluctuations on training expenditures are statistically significant. When training was defined as percent of workforce trained, the aggregate and sectoral channels had similar magnitudes (and of course opposite signs); here, the dollar impact is much larger on the sectoral variable than on the aggregate output one. Overall, the story coming from the continuous definitions of training is similar to the story on training incidence. Firms will train more if the sectors they operate in are hit by relatively favourable idiosyncratic shocks, and will train less when GDP moves above the trend.

3.3 Relaxing strict exogeneity: quantitative results under sequential exogeneity

Strict exogeneity is a strong assumption that does not need to hold when feedback effects are present. While we do control for aggregate and sectoral shocks explicitly, as well as for firm fixed effects, there remains the possibility that a particular firm experiences an unanticipated shock specific only to the firm, and not to the sector or to the economy. To account for this possibility, we relax the strict exogeneity assumption by replacing it with the weaker sequential exogeneity assumption:

$$E(u_{it}\,|X_{i1}, X_{i2}, \ldots, X_{it}, \alpha_i) = 0, \qquad t = 1, \ldots, T. \tag{4}$$

In this formulation, feedback shocks from u_{it} to $X_{i,t+k}$ are accommodated as u_{it} needs to be uncorrelated with all past and current realizations of the endogenous variable X, but not with future ones.

Building on the methodology by Arellano and Bond (1991) and Arellano and Bover (1995), β can be estimated consistently with a generalized method of moments (GMM) estimator. The following moment conditions become available from the first differences equation 2:

$$E(X_{i,t-s}\Delta u_{it}) = 0 \text{ for } s \succeq k. \tag{5}$$

The moment conditions imply that lagged levels of the variables in X from period $t-k$ and earlier can be used as instruments for the differences in period t where the value of k depends on the structure of the error term. If we assume that u_{it} follows an AR(1) process, then $k = 2$ and we are able to use lagged levels of the variables from two periods and earlier on as instruments for the differences in period t.

However, empirical studies and simulations show that the correlation between lagged levels and first differences could be relatively low in certain cases such as highly persistent data, leading to a weak instruments problem and resulting in poor estimates of the differenced GMM estimator.[12] Blundell and Bond (1998) show how further additional moment conditions can be obtained by imposing certain restrictions on the initial moment conditions:[13]:

$$E[\Delta X_{i,t-s}(\alpha_i + u_{it})] = 0 \text{ for } s \succeq 2. \tag{6}$$

Equation (6) implies that we can use the lagged differences of the regressors as instruments in the level equation (1). Therefore, the information provided by the moment conditions (5) and (6) allow us to use the lagged values of levels of the regressors as instruments in the difference equation (2), and the lagged values of differences of regressors as instruments in the level equation 1. This approach results in a system of equations in a GMM framework which leads to considerable efficiency gains as shown by Blundell and Bond (1998, 2000).

The results from this estimation are in Table 6. While quantitatively the magnitude of the coefficients decreases, the sign of the adjustments stays the same. The impacts of the sectoral shock are somewhat larger, and statistically significant, for both specifications (percentage of workforce trained and training expenditures per worker). The impact of GDP fluctuations is negative, though not significant. The technology adoption factor is significant in the first specification (percentage workforce trained) but not in the second (training expenditures).

As far as specification tests go, we look at the validity of the AR(1) assumption as well as the validity of instruments with an IV overidentification test (as by construction we have many instruments). Using the Arellano and Bond serial correlation test, we do not reject the AR(1) structure for u_{it} while we do reject AR(2) for both specifications. This is good news as the validity of the instruments depends on the absence of serial correlation in ϵ_{it} in the AR(1) process $u_{it} = \rho u_{it-1} + \epsilon_{it}$. The lag instruments are for the most part not significant, except for the lag dependent variable. This is problematic as it can be indicative of weak instruments; nevertheless, the instruments are jointly significant. Furthermore, given the large number of

[12] See Blundell and Bond (1998, 2000) for details.

[13] Joint stationarity of y_{it} and X_{it} as a sufficient but not necessary condition.

instruments, we also report the Hansen test of over identifying restrictions, which is consistent in the presence of non-spherical errors. In both specifications the tests show that instruments satisfy the orthogonality conditions implied by the moment conditions.

The main message from this sensitivity analysis is that, even under a weaker identification assumption, the sectoral channel still has an important effect: the better the relative position of a sector, the more likely are the firms in this sector to train. The aggregate channel is still present even though less strongly: if the whole economy is experiencing a positive shock, all firms will have incentives to train less.

3.4 Further evidence on training and sectoral reallocation

We first bring suggestive evidence that the idiosyncratic sectoral shock channel can be associated with the reallocation of lower-skill workers from the sectors doing relatively poorly to the sectors doing relatively better. Table 7 reports the correlations between the skill distribution of new hires and the idiosyncratic sectoral shock.[14] There is a positive correlation between the sectoral shock and the fraction of new hires that are unskilled "production" workers, while all other correlations are negative. (Note that there is no correlation between the skill group of the new hires and the aggregate business cycle.) These results seem to indicate that sectors hit by relatively better idiosyncratic shocks are more likely to increase their unskilled workforce.

Moreover, we use the worker side of the WES panel to investigate what factors influence the training decision of workers. To this extent we estimate a logit fixed effects model separately for men and for women controlling for the usual determinants of training: age, marital status, education, occupation, firm size, tenure, and interactions between tenure and education to allow for heterogeneous impacts of tenure across education groups.[15] While we do not know the firm or sector where a new worker is coming from (nor do we follow workers leaving a firm), we do know whether a new worker with job tenure less than one year is also new to the sector. This is our main control variable; as seen from Table 8, its impact is large and positive. Put differently, a worker who is new to the firm and new to the sector is more likely to get trained, all else being equal. This effect seems large, even more so for men than for women, and it indicates

[14] Workers are only interviewed if they stay within the firm; thus, we cannot follow them as they move across firms or sectors.

[15] While fixed effects will wipe out the effect of non-varying regressors, fixed effects are still paramount in our estimation, as they mitigate to a very large extent issues of self-selection into training.

that new entrants into a sector are more likely to get trained.

4 A Framework for Interpreting the Results

To illustrate the relationship between output shocks - aggregate and sectoral - and training by firms, we present here a theoretical model based on the Pissarides (2000) matching model with heterogeneous workers. In interpreting our results we propose a theoretical framework consisting of a straightforward extension of the Pissarides model to include training decisions by firms. While here we give only the intuition of the mechanisms and equilibrium predictions of the model, a more technical exposition and equations is presented in Appendix I.

Since we directly observe and control for innovation and the adoption of new technologies, what we measure in the empirical section is the direct effect of output fluctuations on training beyond innovation. The negative coefficient on the aggregate output channel conforms to the literature arguing that more training occurs during downturns when the opportunity cost of training is lower. We conjecture that the sectoral impact on training comes from a reallocation mechanism: sectors doing relatively better will attract workers from sectors doing relatively worse. The new workers will need skills training in the new sector. Moreover, it is more likely that workers who move are low-skill workers with less investment in sector-specific human capital (which gets destroyed upon a move).

Consider the basic Pissarides matching model in continuous time. Unemployed workers search for a job, and firms post vacancies. Once a worker meets a vacancy, a match-specific productivity α is observed, and a match is formed if and only if the realized productivity is above a given threshold R. In every period, the productive match is subject to a productivity shock y. We introduce aggregate and sectoral fluctuations via this productivity shock as $y = p\epsilon$, with p representing the aggregate shock and ϵ the idiosyncratic sectoral one. Only matches with ongoing productivity above the reservation threshold R are preserved.

In our extension of the model, training is made available to workers at some cost for the firm, resulting in increased worker productivity according to a specified human capital function. Training will take place as long as the benefit from training is higher than the cost. In equilibrium, this implies that two categories of workers will be engaged in productive matches: (i) workers with high ex-ante productivity, who are above the reservation threshold R even without training, and (ii) a second group of workers with medium to low ex-ante productivity; without training these workers would fall below the reservation productivity threshold R, while with training their productivity can jump above the reservation threshold. This implies that the reservation threshold R gets de facto expanded into a productivity interval $[R_0, R]$. Workers falling in this interval are still able to form productive matches if they get training, while workers with ex-ante productivities above R form productive matches without training. There is a third category of workers, those with the lowest productivity (below R_0) who would not meet the reservation threshold R, even with training. Consequently, they will not get trained and will not engage in a productive match.

When an aggregate negative shock p hits all sectors, training will increase as long as the marginal cost of training is higher than the marginal benefit with respect to p, which is easy to achieve under very reasonable parameterizations of the human capital function. This generates the expected counter-cyclical channel of more training during periods of output decline.

When an idiosyncratic shock ϵ hits the sectors, there will be worker reallocation from the low-shock to the high-shock sectors. The workers with smaller present values in the low-shock sectors are likely to leave; they are also lower-productivity workers. (Some but not all of them may have been below the employment productivity threshold R_0 in the sector they leave). The high-shock sectors will get an influx of lower-productivity workers, both decreasing the average match quality α and increasing worker congestion in the high-shock sectors. As a result, more matches will fall within the training productivity interval, and thus more training will take place in the relatively better sectors.

5 Conclusion

The sectoral analysis has been shown to be very important when specifying the links between the aggregate business cycle, sectoral idiosyncratic shocks, firm innovation, and the incidence and intensity of training. We find training to be counter-cyclical – firms train more in downturns, while sectoral shocks have a positive impact on training incidence – more training when the sector is in a relatively better position. As the magnitudes of these adjustments are of similar order, we should not ignore either one.

We believe the finding of two opposing channels through which output fluctuations affect training decisions is relevant for at least three reasons: (i) it gives us better insights into firms' training decisions over the business cycle, (ii) it quantifies how aggregate and sectoral shocks play into the human capital accumulation decision, and (iii) it helps policy-makers understand that fluctuations in training incidence may be an optimal response to macroeconomic shocks, and not necessarily an indicator of underinvestment in training.

The previous literature has argued that there is a role for government intervention in employer training. In deciding how much training to provide, firms will take into consideration how likely the workers are to stay with the firm once the training is completed. If private returns to training are large but firms do not train for fear of losing workers to higher-paying jobs, then it is socially optimal for governments to intervene by providing training directly (or through incentives for workers and firms to increase training). Nevertheless, we point out that governments should exercise caution when interpreting statistics on the incidence of training by firms. The aggregate and sectoral output fluctuations documented here can inform policy-makers whether observed trends in training are healthy, as dictated by economic circumstances, or whether firms under-invest in training and that therefore government intervention should be recommended.

Finally, from a theoretical standpoint, we highlight the importance for any models of firm training to incorporate mechanisms coming from both aggregate and sectoral output fluctuations. Such models help us get a better understanding of the training decisions by firms. Moreover, any model used for policy prescriptions regarding training needs to be consistent with both aggregate and idiosyncratic facts; a model consistent with only one or the other will provide wrong policy recommendations.

References

Arellano, M., and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277–297.

Arellano, M., and O. Bover (1995): "Another Look at the Instrumental Variables Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29–51.

Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 11–143.

——— (2000): "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19(3), 321–340.

Dearden, L., H. Reed, and J. V. Reenen (2006): "The Impact of Training on Productivity and Wages: Evidence from British Panel Data," *Oxford Bulletin of Economics and Statistics*, 68(4), 397–421.

deJong, D., and B. Ingram (2001): "The Cyclical Behavior of Skill Acquisition," *Review of Economic Dynamics*, (4), 536–561.

Dellas, H., and P. Sakellaris (2003): "On the Cyclicality of Schooling: Theory and Evidence," *Oxford Economic Papers*, (55), 148–172.

Devereux, P. (2000): "Task Assignment over the Business Cycle," *Journal of Labor Economics*, 18(1), 98–124.

King, I., and A. Sweetman (2002): "Procyclical Skill Retooling and Equilibrium Search," *Review of Economic Dynamics*, 5(3), 704–717.

Lynch, L., and S. Black (1998): "Beyond the Incidence of Training: Employer-Provided Training," *Industrial and Labour Relations Review*, 52(1), 64–81.

Pissarides, C. (2000): *Equilibrium Unemployment Theory*. Cambridge, MA: MIT Press.

Turcotte, J., A. Leonard, and C. Montmarquette (2003): "New Evidence on the Determinants of Training in Canadian Business Locations.," *The Evolving Workplace Series*, Statistics Canada and HRDC, cat. no. 71-584-MIE.

Wooldridge, J. (2005): "Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, (20), 39–54.

Figure 1: Log GDP Deviations from Trend: 1982-2007. Scale in percentages.

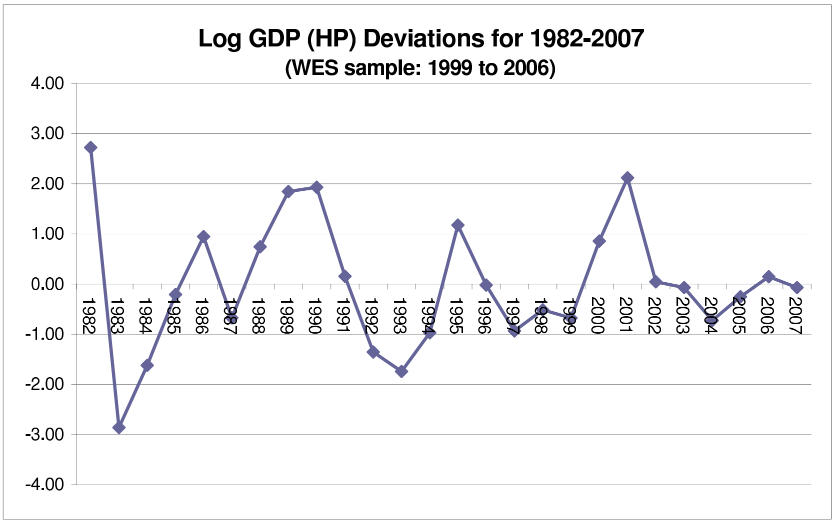

Table 1: Statistics for Training Incidence and for Firm-Specific Variables

Variable	Description	Mean	Std. Dev.
Classroom Training Indicator		0.388	0.487
% Workforce Trained		22.3%	34.75
Training Expenditures per worker		$178.6	$502.8
Firm size	Number of workers employed by the workplace	19.28	48.37
Innovation	Adoption new technology/innovation by the workplace	0.493	0.499
The most dominant sales market of the firm			
	Market: Local	0.850	0.357
	Market: Canada	0.100	0.299
	Market: World	0.050	0.218
Unionized	Indicator whether workplace is unionized	0.058	0.235
Multiple loc.	Indicator whether workplace belongs to multiple-location firm	0.550	0.497
% of workforce in each skill group			
	Staff: % Administrative	0.199	0.273
	Staff: % Managerial	0.202	0.213
	Staff: % Other	0.061	0.196
	Staff: % Professionals	0.062	0.172
	Staff: % Sales	0.122	0.240
	Staff: % Production	0.159	0.265
	Staff: % Technical	0.195	0.306

Source: WES 1999-2006, firm side (2263 firms). With slight abuse of notation, we use the terms "firm" and "workplace" interchangeably to denote a plant.

Table 2: Sectors in the Analysis

Sector	Relative Size (%)
Forestry and Mining	4 %
Construction	11 %
Transportation, Warehouse, Wholesale Trade	15 %
Information, Communication and Utilities	8 %
Finance and Insurance	9 %
Real Estate	5 %
Business Services	9 %
Education and Health	4 %
Manufacturing	24 %
Retail Trade and Consumer Services	11 %

Source: WES 1999-2006, firm side (2263 firms). Relative size of each sector expressed as percentage of firms in the sector relative to total firms.

Table 3: The Impact of Aggregate and Sectoral Output Fluctuations on Training Incidence: Extensive Margin

	CT Indicator		CT and OJT Indicator	
	Coef.[a]	Std. Err.	Coef.[a]	Std. Err.
	(1)	(2)	(3)	(4)
GDP fluctuations [b]	-0.072	0.003	-0.040	0.003
Sector to GDP ratio (in %)	0.034	0.002	0.022	0.003
Innovation	0.629	0.005	0.643	0.005
Market: Canada [c]	0.406	0.011	0.580	0.011
Market: World	0.449	0.020	-0.491	0.019
ln (Firm size)	0.423	0.008	0.484	0.008
Multiple locations	0.112	0.006	-0.132	0.007
Unionized	0.133	0.020	-0.329	0.020
Staff: % Administrative [c]	0.260	0.020	0.154	0.019
Staff: % Managerial	0.560	0.020	0.610	0.019
Staff: % Other	1.023	0.021	0.265	0.020
Staff: % Sales	0.634	0.021	0.669	0.020
Staff: % Production	0.734	0.019	0.538	0.019
Staff: % Technical	-0.083	0.018	0.162	0.019
GDP Trend	0.001	0.000	0.0002	0.0000

Dependent variable is an indicator whether the firm has trained or not. Data from WES 1999-2006, Firm side. The number of observations (firms) is 8,913 (1,120). Estimation using fixed-effects logit.

[a] Coefficients are not marginal effects, but the ratio of any two coefficients is the same as the ratio of the two marginal effects.

[b] HP detrended log GDP.

[c] Base category: Local output market; % Professionals in workforce.

Table 4: The Impact of Aggregate and Sectoral Output Fluctuations on Training Decisions: All Firms.

	% Workforce Trained			Training Expenditures per Worker		
	Coef.	Std.Err.	Robust Std.Err.	Coef.	Std.Err.	Robust Std.Err.
	(1)	(2)	(3)	(4)	(5)	(6)
GDP fluctuations [a]	-0.654	0.241	0.269	-4.286	3.597	8.229
Sector to GDP ratio (in %)	0.530	0.228	0.179	14.509	3.406	3.162
Innovation	5.979	0.509	1.783	53.208	7.604	25.493
Market: Canada [b]	2.498	1.008	2.203	50.423	15.060	25.950
Market: World	5.431	1.871	4.184	-0.066	27.943	29.910
ln (Firm size)	-0.142	0.773	1.700	-57.999	11.547	38.391
Multiple locations	0.271	0.561	2.070	10.935	8.382	28.015
Unionized	4.865	1.833	4.582	55.636	27.379	33.687
Staff: % Administrative [b]	2.506	1.838	3.068	-37.974	27.448	76.526
Staff: % Managerial	2.806	1.889	2.859	55.753	28.223	69.274
Staff: % Other	4.459	1.979	4.480	35.480	29.563	76.312
Staff: % Sales	7.394	1.949	4.364	31.671	29.119	82.171
Staff: % Production	3.663	1.836	5.081	14.117	27.424	82.691
Staff: % Technical	-2.222	1.836	4.949	-22.862	27.426	116.969
GDP Trend	0.008	0.003	0.010	0.091	0.049	0.154
Constant	3.562	4.235	11.121	37.748	63.260	155.615

Data from WES 1999-2006, Firm side. The number of observations (firms) is 16,208 (2,037). Estimation using fixed-effects OLS.

[a] HP detrended log GDP.

[b] Base category: Local output market; % Professionals in workforce.

Table 5: The Impact of Aggregate and Sectoral Output Fluctuations on Training: Intensive Margin. Sample Restricted to Firms who Train.

	% Workforce Trained			Training Expenditures per Worker		
	Coef.	Std.Err.	Robust Std.Err.	Coef.	Std.Err.	Robust Std.Err.
	(1)	(2)	(3)	(4)	(5)	(6)
GDP fluctuations [a]	-0.895	0.298	0.336	-6.995	4.609	12.148
Sector to GDP ratio (in %)	0.703	0.280	0.283	19.040	4.285	4.296
Innovation	8.124	0.631	2.193	75.903	9.777	34.584
Market: Canada [b]	3.363	1.201	2.987	72.632	18.149	30.782
Market: World	8.727	2.291	5.520	10.860	34.553	33.021
ln (Firm size)	-0.166	0.949	1.944	-84.155	14.489	49.956
Multiple locations	0.584	0.681	2.643	19.048	10.507	35.741
Unionized	5.769	2.126	5.315	58.835	32.087	35.016
Staff: % Administrative [b]	2.901	2.248	4.768	-87.862	36.223	105.465
Staff: % Managerial	3.994	2.369	4.302	119.290	36.987	91.543
Staff: % Other	6.136	2.458	6.611	45.598	38.303	102.191
Staff: % Sales	10.074	2.396	6.556	29.641	37.879	116.107
Staff: % Production	4.985	2.243	7.687	14.900	35.046	108.943
Staff: % Technical	-2.026	2.151	5.919	-29.940	33.448	151.289
GDP Trend	0.010	0.004	0.012	0.129	0.061	0.187
Constant	5.849	5.126	13.991	97.866	79.372	194.019

Data from WES 1999-2006, Firm side. The sample is restricted to firms who train. The number of observations (firms) is 14,198 (1,784). Estimation using fixed-effects OLS.

[a] HP detrended log GDP.

[b] Base category: Local output market; % Professionals in workforce.

Table 6: The Impact of Aggregate and Sectoral Output Fluctuations on Training: Instrumenting for Potential Endogeneity

	% Workforce Trained		Training Expenditures	
	Coef.	Std. Err.	Coef.	Std. Err.
GDP fluctuations	-0.85	1.69	-27.26	39.26
lag	-0.36	0.86	-18.20	23.00
Sector to GDP ratio (in %)	2.35	1.16	32.99	14.92
lag	-0.92	0.70	10.80	42.24
Training (LHS) lag	0.21	0.04	0.42	0.15
Innovation	9.08	2.04	104.30	86.05
lag	0.04	1.60	-28.66	61.14
ln (Firm size)	-3.43	2.85	-291.18	657.84
lag	-0.23	2.84	-93.09	131.24
Staff: % Administrative	0.57	9.00	-384.31	471.67
lag	-4.13	8.09	-994.44	363.89
Staff: % Managerial	7.63	8.93	154.24	315.15
lag	11.72	9.46	-287.94	300.75
Staff: % Other	0.99	9.30	29.33	301.11
lag	7.83	8.71	-582.62	317.15
Staff: % Sales	6.97	10.28	337.00	299.49
lag	20.30	8.83	-148.01	312.99
Staff:% Production	-0.22	9.22	138.02	274.34
lag	0.33	7.81	-499.83	316.58
Staff: % Technical	-6.97	9.54	-46.07	335.82
lag	10.02	8.95	-512.12	413.94
Multiple locations	-5.86	3.63	166.18	57.59
lag	2.45	2.58	24.00	55.23
Unionized	7.03	5.88	9.00	73.35
lag	1.38	5.22	-91.16	75.76
Market: Canada	0.65	3.31	87.59	72.28
lag	-0.91	2.75	-42.90	96.34
Market: World	6.30	5.58	-30.77	75.50
lag	-3.52	5.22	-47.12	88.88
GDP Trend	-0.22	0.73	-9.21	25.12
lag	0.22	0.72	10.99	24.62
Constant	23.47	46.93		
AR(1) test (Arrelano-Bond)	z=-11.99	P>z=0.000	z=-2.47	P>z=0.013
AR(2) test (Arrelano-Bond)	z=0.18	P>z=0.861	z=1.20	P>z=0.229
Overid. test (Hansen)	$\chi^2(45)$=91.10	$P > \chi^2$=0.099	$\chi^2(45)$=61.01	$P > \chi^2$=0.056

[a] Data from WES 1999-2006, Firm side. 12,642 observations (1818 firms).
[b] HP detrended log GDP.
[c] Base category Local output market; % Professionals in workforce.

Table 7: Correlations between the Skill Category of New Hires and the Sectoral and Aggregate Fluctuations

Skill category	Correlation with Sector to GDP ratio	Correlation with HP detrended log(GDP)
Administrative	-0.148*	0.009
Sales	-0.091*	-0.023
Managerial	-0.019	-0.007
Professional	-0.090*	-0.003
Technical	-0.066*	-0.003
Production	0.301*	0.011
Other	-0.068*	0.004

* Correlations significant at 1% level.
Data from WES 1999-2006, Firm side.

Table 8: Determinants of Training Incidence, Worker Side

Variables	Coefficients Men	Std. Err.	Coefficients Women	Std. Err.
Worker New in Sector	0.431	0.004	0.201	0.004
Age	-0.193	0.001	-0.204	0.001
Married	0.271	0.004	0.041	0.004
Community College	0.394	0.006	-0.185	0.006
University	0.505	0.010	0.049	0.010
Post-graduate	-0.212	0.016	-1.621	0.019
Unionized	0.503	0.004	0.204	0.004
Occ: Managerial	0.327	0.007	0.342	0.007
Occ: Technician	-0.033	0.006	-0.242	0.005
Occ: Administrative	0.008	0.009	-0.581	0.005
Occ: Sales	-0.856	0.012	-1.245	0.008
Occ: Production	-0.036	0.008	0.444	0.008
Job tenure	-0.040	0.001	-0.067	0.001
Job tenure ^{2}	-0.001	0.000	0.001	0.000
Tenure * College	0.041	0.001	0.045	0.001
Tenure * University	-0.008	0.001	0.030	0.001
Tenure * Postgrad	0.031	0.002	0.036	0.001
Small firm	0.057	0.005	1.021	0.006
Medium firm	-0.016	0.007	1.417	0.008
Large firm	-0.252	0.010	1.480	0.011

Data from WES 1999-2006, Worker side. The number of observations (workers) is 18,659 (8,697) for men and 14,717 (6,749) for women. Estimation using fixed-effects logit.

Appendix I Sketch of Mortensen-Pissarides model with training

Firms open vacancies whenever they want to fill a job. Keeping a vacancy open implies a cost c. The rate at which unemployed workers and open vacancies meet, in each sector is regulated by meeting functions $m(v_i, u)$ that depends on the number of unemployed workers and vacancies created in the particular sector i. Once there is a meeting firms observe the worker specific productivity α and decide if the candidate is suitable for the job. A productive match is formed if α is above the reservation value R_i. Upon creating a match the firm evaluates the opportunity to train the worker. Depending on the productivity of the worker and whether the worker has been trained or not a wage $w^k(\alpha)$ is paid, with $k = u, \tau$ for untrained and trained respectively. Training, as well as the productivity α, are specific to the match: if the match is dissolved the worker returns to the pool of unemployed workers with unknown productivity, and with the same expected productivity she had before the match (and same as everybody else in that pool). After a match is created shocks can arrive at a rate λ which will dissolve the match and let the worker be unemployed again. Wages are set by Nash bargaining.

Following Pissarides (2000) the meeting function is written as $m(v_i, u) = m(1, \frac{u}{v_i})v_i \equiv q(\theta_i)v_i$, where $\theta_i = \frac{v_i}{u}$ is market tightness in sector i. Given the meeting function the ratio at which vacancies are filled can be defined as $q_i^f = q(\theta_i)\int_{R_i}^{b} dF(\alpha)$, where b is the upper limit of the shock distribution and r is the reservation value. The ratio at which unemployed workers find a job is respectively given by: $q_i^w = q(\theta_i)\theta_i \int_{R_i}^{b} dF(\alpha)$.

Appendix I.1 Value of a match to an employer

The value of a job to an employer depends on the productivity specific to that match and the level of training given to the worker. We assume that the level of training is decided at the starting of a match (since empirical evidence suggests that most training takes place early in the employer tenure), and that the cost of training is paid by the employer every period the worker is employed (such an insurance-type cost scheme enables firm training even if training is in transferable general skills). Output is the product of the shock α and the productivity parameter $y_i = p\epsilon_i$, where p is an aggregate productivity shock and ϵ_i a sector-specific idiosyncratic one.

The value of a match to an employer who trains $J_i^\tau(\alpha)$ − or not $J_i^u(\alpha)$ − is given by:

$$rJ_i^u(\alpha) = y_i\alpha - w_i^u(\alpha) - \lambda J_i^u(\alpha). \tag{7}$$

$$rJ_i^\tau(\alpha) = y_ih(\alpha) - w_i^\tau(\alpha) - C(y\alpha) - \lambda J_i^\tau(\alpha). \tag{8}$$

Here $h(\alpha)$ is a function that describes how productivity increases with training and it is assumed to be increasing in α, $C(y\alpha)$ is the cost of training, and w^τ and w^u the wage rates offered to the trained and respectively untrained workers.[16]

The asset equations above describe the value of a match. Training is required for workers with productivity levels above the reservation threshold R_i but below $\alpha_{\tau,i}$, while is not required for workers with higher productivity. The asset equation that describes the expected value of a match before the meeting takes place is given by:

$$J_i^e = \int_{R_i}^{\alpha_{\tau,i}} J_i^u dF(v) + \int_{\alpha_{\tau,i}}^{b} J_i^\tau(v)dF(v). \tag{9}$$

where the superscript e indicates the expectation conditional on α being greater than the productivity threshold R_i.

Appendix I.2 Value of a match to a worker

The value of a match to a worker is determined by the following asset equations:

$$rW_i^u(\alpha) = w_i^u(\alpha) + \lambda[U - W_i^u(\alpha)], \tag{10}$$

$$rW_i^\tau(\alpha) = w_i^\tau(\alpha) + \lambda[U - W_i^\tau(\alpha)]. \tag{11}$$

Similarly to the previous case for the employer we can write,

$$W_i^e = \int_{R_i}^{\alpha_{\tau,i}} W_i^\tau(v)dF(v) + \int_{\alpha_{\tau,i}}^{b} W_i^u(v)dF(v). \tag{12}$$

[16]Note that if we were interested in say the optimal amount of training T offered, we could introduce it via the benefit and cost of training $h(\alpha, T)$ and $C(T, y\alpha)$ where T can denote the amount of training, $h(\alpha, T)$ is concave in T and $C(T, y\alpha)$ is convex in T. Here we focus only on the extensive margin instead.

Appendix I.3 Value of a vacancy and of unemployment

The value of setting a vacancy to an employer is

$$rV_i = -c + q_i^f[J^e - V],$$

where c is the vacancy posting cost. In equilibrium free entry sets the value of a vacancy to zero.

The value of unemployment to a worker depends on the number and the conditions of all sectors in the economy, since each unemployed worker can be matched stochastically with any of the firms opening vacancies in each sector. For simplicity, we assume that there are only two sectors in the economy indexed by $i = 1, 2$. In this case we have that the total number of vacancies formed in the economy is given by $v = v_1 + v_2$ and the overall tightness of the economy is described by $\theta = \theta_1 + \theta_2$. The value of unemployment is then

$$rU = z + q_1^w[W_1^e - U] + q_2^w[W_2^e - U],$$

where z represents unemployment contingent income.

Appendix I.4 Wages and Training

Assuming that the wage rates are set following the Nash bargaining rule with bargaining power β, after some algebra we can derive the wage rate for trained and untrained workers,

$$w_i^\tau(\alpha) = \beta[y_i h(\alpha) - C(y_i\alpha)] + (1-\beta)z + \beta c\theta \tag{13}$$

$$w_i^u(\alpha) = \beta y_i\alpha + (1-\beta)z + \beta c\theta. \tag{14}$$

Reservation value for training

Training occurs as long as $J^\tau(\alpha) \geq 0$ and up to the point where the value of an untrained match is equal to the value of a trained match, that is, $J^u(\alpha_\tau) = J^\tau(\alpha_\tau)$, or:

$$y_i[\alpha_{\tau,i} - h(\alpha_{\tau,i})] = w_i^u(\alpha_{\tau,i}) - w_i^\tau(\alpha_{\tau,i}) - C(y_i\alpha_{\tau,i})$$

$$y_i[h(\alpha_{\tau,i}) - \alpha_{\tau,i}] = C(y_i\alpha_{\tau,i}) \tag{15}$$

Reservation value for hiring

The reservation value for hiring is set by the following equation $\max\{J^u(R), J^\tau(R)\} = 0$. Notice that, as long as $R < \alpha_\tau$ (and therefore some training occurs), the relevant condition can be re-written as $J^\tau(R) = 0$, or $(1-\beta)[yh(R) - C(yR)] = (1-\beta)z + \beta c\theta$.

$$yh(R) - C(yR) = z + \frac{\beta}{1-\beta} c\theta \tag{16}$$

Appendix I.5 Aggregate Shocks

Assume that sectors 1 and 2 are identical (because the assumption that sectors are identical we drop the subscript "i"), and focus on how the aggregate productivity shock p influences the decision to train. When p changes the two reservation productivity thresholds α_τ (for training decisions) and R (for hiring decisions) may also change.

From equation (15) we can see that if and how α_τ changes depends on what we assume about the functions h and C. We can therefore find appropriate functions that deliver the predictions we observe from the data. In particular, if we assume that

$$[h(\alpha) - \alpha] < \frac{\partial C(y\alpha)}{\partial y} \tag{17}$$

when overall productivity decreases, training gets more convenient because its cost decreases faster than the relative benefit and α_τ raises.

When y decreases θ should decrease as well since unemployment increases for the whole economy more than vacancies do. Therefore, the RHS of (16) decreases and the so LHS has to decrease as well. If, like in the basic Pissarides model with stochastic job matching, R should also increase, then the higher R in this case might imply lower training because relatively more workers with higher productivity and no need for training are going to be hired. (Note there is no such problem if R does not increase, or if at the same time α_τ raises sufficiently). The final effect would depend on the parametrization of the model and the choice of h and C, and as such, we can always find reasonable functions for which the first channel on α_τ prevails and generates counter-cyclical training.

Appendix I.6 Sectoral reallocation

The impact of the idiosyncratic shock ϵ_i is easier to show when we think of the adjustments that happen when one sector only, say for instance sector 2, experiences a negative shock, that is, ϵ_2 is lower. Re-proposing equation (16) for sector 1 we have,

$$y_1 h(R) - C(y_1 R) = z + \frac{\beta}{1-\beta} c\theta \tag{18}$$

The new steady state is characterized by higher unemployment and lower θ, adjustments through which sector 2 influences sector 1. In equation (18) the RHS is lower, and, because y_1 does not go down, the only way to re-establish the equality is by reducing R. (Also note that with no change in ϵ_1, $\alpha_{\tau,1}$ does not change, as θ does not enter in its determination.) Therefore, because the pool of workers to be trained is now larger, training will increase in sector 1. In other words, the outside option for employees working in sector 1 looks worse after sector 2 is hit by the shock. This affects their salaries determined by the Nash bargaining through the lower value of unemployment. However, sector 1 still has the same productivity as before, therefore for lower wages the firms will be able to hire more workers. Firms do that by lowering the reservation value of the productivity shock R_1. That is, more low skill workers will be hired in sector 1, and they are workers who are more likely to need training.

Appendix II Firm-specific determinants of training incidence

Table II1: The Impact of Aggregate and Sectoral Output Fluctuations on Training Incidence: Extensive Margin. OLS analysis (dychotomous dependent variable).

	CT Indicator Coef.	Std. Err.[a]	Robust[a] Std. Err.
	(1)	(2)	(3)
GDP fluctuations [b]	-0.010	0.004	0.005
Sector to GDP ratio (in %)	0.005	0.002	0.001
Innovation	0.085	0.020	0.068
Market: Canada [c]	0.053	0.049	0.108
Market: World	0.042	0.071	0.159
ln (Firm size)	0.061	0.023	0.051
Multiple locations	0.014	0.026	0.095
Unionized	0.013	0.037	0.093
Staff: % Administrative [c]	0.053	0.039	0.065
Staff: % Managerial	0.085	0.046	0.069
Staff: % Other	0.135	0.068	0.153
Staff: % Sales	0.092	0.047	0.105
Staff: % Production	0.097	0.087	0.242
Staff: % Technical	-0.029	0.039	0.104
GDP Trend	0.000	0.000	0.000
Constant	-0.073	0.149	0.390

Data from WES 1999-2006, Firm side. The number of observations (firms) is 8,913 (1,120). Estimation using fixed effects OLS.

[a] Standard errors are computed in two ways: without accounting for clustering by sector in column (2), and accounting for clustering in column (3). Section 3.2 has more details on the difference in approaches and its substantive implications.

[b] HP detrended log GDP.

[c] Base category Local output market; % Professionals in workforce.