The B.E. Journal of Macroeconomics

Advances

Volume 10, *Issue* 1 2010 *Article* 28

The Impact of Aggregate and Sectoral Fluctuations on Training Decisions

Vincenzo Caponi* Burc Kayahan†
Miana Plesca‡

*Ryerson University, Institute for the Study of Labor, and The Rimini Center for Economic Analysis, vcaponi@ryerson.ca

†Acadia University, ckayahan@acadiau.ca

‡University of Guelph, miplesca@uoguelph.ca

Recommended Citation

Vincenzo Caponi, Burc Kayahan, and Miana Plesca (2010) "The Impact of Aggregate and Sectoral Fluctuations on Training Decisions," *The B.E. Journal of Macroeconomics*: Vol. 10: Iss. 1 (Advances), Article 28.

Available at: http://www.bepress.com/bejm/vol10/iss1/art28

The Impact of Aggregate and Sectoral Fluctuations on Training Decisions*

Vincenzo Caponi, Burc Kayahan, and Miana Plesca

Abstract

The literature on training has pointed out that macroeconomic fluctuations can have a positive or a negative effect on training decisions. 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. The first contribution of this paper is to document, using the Canadian panel of Workplace and Employee Survey (WES), that (i) training moves counter-cyclically with 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 sectors doing relatively better). This second fact is novel and unexplored. Overall, the results show that the firms' decisions to train are quite complex; in order to fully understand them, one needs to take into account not only the change in aggregates, but also the relative position of each sector in the economy. The second contribution of the paper is to illustrate the mechanisms at work by incorporating training decisions into a standard Mortensen-Pissarides model. In the standard model, production takes place if workers' productivity is above a reservation threshold. In our extension, this threshold gets expanded into a whole interval within which production takes place if workers are trained. The quantitative analysis from the calibrated model illustrates the counter-cyclical opportunity cost adjustment from aggregate shocks and the pro-cyclical adjustment coming from sectoral reallocation.

KEYWORDS: training, human capital, business cycle, sectoral shocks

*This research was financially supported by a grant from the Canadian Labour Research Network (CLSRN). We thank Gueorgui Kambourov, Louise Grogan, Chris Robinson, the editor, two anonymous referees, and participants at the 2010 Society for Economic Dynamics conference, the 2010 Rimini Conference in Economics and Finance, the 2009 Canadian Economics Association meeting, the 2008 CLSRN Workshop, the 2008 Atlantic Canada Economic meeting, and seminars at the University of Guelph, Wilfrid Laurier University, and Mount Allison University for very helpful comments and discussions.

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 an establishment'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 establishments hoard labor 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.

This paper has two main contributions. First, using the Canadian Workplace and Employee Survey (WES), we find that (i) training moves counter-cyclically with 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). Second, in order to illustrate the mechanisms at work, we incorporate training decisions into a standard Mortensen-Pissarides model. The quantitative analysis from the calibrated model illustrates the magnitude of the two training channels: opportunity cost of training, which is low when productivity is low, and reallocation, which is high when a sector's position improves.

In order to measure the quantitative impact of aggregate and sectoral output fluctuations on training incidence, we use the Canadian WES dataset, an eight-year panel of workplaces from 1999 to 2006, representative of all industries except for agriculture and public administration. The unit of analysis is the workplace, or establishment. The WES is a very appealing data set because response rates are consistently high across all panel years, and sample sizes are relatively generous, especially compared to other establishment-level data. Most importantly, the panel nature of the WES allows us to remove the unobserved firm-specific fixed effects in the empirical analysis.[1]

For employer training we focus on two main definitions: (i) "the extensive margin of training", a binary indicator whether the establishment has provided training or not, and (ii) "the intensive margin of training", a continuous measure, conditional on the establishment providing some training, expressed either as a percentage of the workforce trained or as training expenditures per worker. We measure the fluctuations in aggregate output as deviations of log total output from a Hodrick-Prescott (HP) trend, while the relative position of a sector is measured as the share of that sector in total output. In particular, we think of the *aggregate* output fluctuations as fluctuations of output around a (mostly-upward) trend, and identify them from the time-series dimension of the aggregate data. We think of the *idiosyncratic* sectoral output fluctuations as the share of a particular sector in total output, and measure them from the cross-sectional position of sectors in the economy.[2] Our major findings are that (i) training moves counter-cyclically with 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 one percentage point and decreases training expenditures by $7 per worker per year, representing 3.9% of the mean annual training expenditures per worker, or $178.6 (in constant 2000 dollars). A one-percentage point increase in the share of a sector's output, controlling for aggregate shocks, *increases* the probability of training by 0.5 percentage points and increases training expenditures per worker per year by $19 (that is, 10.7%) for the establishments that train. In all specifications, we control for a host of other factors, out of which the most relevant is innovation.[3] Innovation has the expected sign: establishments that innovate more train more.

[1] Moreover, the WES is a linked employer-employee survey. While only very few workers are interviewed per establishment, and are therefore not representative of the establishment's workforce, we can still use information from the worker side as supporting evidence for sectoral reallocation and worker training.

[2] Note that to construct the aggregate output trend and its deviations, we use a longer time series of the aggregate real GDP for the whole economy. For the sectors, we use the WES sample, weighted appropriately using survey weights. A sector's share is its own output divided by the sum of all sectors' output in the sample.

[3] Innovation is self-reported by establishments, and is defined as a binary indicator whether the establishment has innovated or adopted new technologies.

The channels of aggregate and sectoral output fluctuations documented here, however, operate even beyond innovation. Moreover, the results are robust to different specifications, including ones where we instrument for potentially endogenous establishment covariates (such as innovation) using the longitudinal structure of the data to construct the instruments.

Like most of the literature, we contend that the negative impact of aggregate output fluctuations on employer training is due to the opportunity cost of training (foregone output), which is lower in downturns. On the other hand, the fact that establishments from sectors that are doing relatively better have an incentive to train more is novel and unexplored. We contend that there are two mechanisms at work in this case. First, insofar as positive sectoral fluctuations are related to the adoption of new technologies, establishments 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) establishments that innovate or adopt new technologies train more; (ii) there is an increase in training incidence by establishments 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.

Our results show that the employers' decisions to train are quite complex. In order to fully understand them, one needs to take into account not only the change in aggregates but also the relative position of each sector in the economy. For example, throughout an expansion of the economy, some sectors could be expanding much faster than others. As a result, even though training in the economy as a whole would be declining (due to the aggregate effect), training in the fastest expanding sectors would decline the least and could actually increase (due to the offsetting sectoral effect). Similarly, during a recession some sectors could be experiencing a much larger decline than others. Then, even though training in the economy as a whole would be increasing, training in the most affected sectors would increase the least and could actually decline.

In order to isolate and illustrate the mechanisms at work, we extend a Mortensen-Pissarides model by introducing training decisions in this framework. We see the model as a contribution in itself, and also as a tool that can help anchor and motivate our empirical findings. In the model, unemployed workers meet vacancies, and a match is formed if the productivity of the match is above an endogenous threshold. Output is the product of three idiosyncratic productivities: individual (match-specific) productivity, sectoral productivity, and aggregate productivity. Training is available at a cost, with training benefits and costs expressed as functions of output. The decision to train occurs at the beginning of the match, and it depends on a productivity threshold interval for workers who can produce if and only if they are trained. This expands the productivity reservation threshold from the standard Mortensen-Pissarides model into an interval: (i) workers whose productivities fall below the lower end of the training interval do not produce, (ii) workers with productivities within this interval engage in a productive match, but only if they are trained, and (iii) workers with productivities above the higher end of the interval engage in a productive match without training.

In this setting, aggregate and sectoral fluctuations change both ends of the interval threshold for training. As long as the benefit of training increases more slowly than the cost when output increases, a positive output shock, be it aggregate or sectoral, will decrease both reservation productivities − the lower one with training, and the upper one without training − shifting the training interval to the left. The productivity threshold without training goes down in both cases, with the aggregate and with the sectoral shock, but it goes down by much more when the sectoral shock hits because of worker reallocation. Following a positive sectoral shock, workers flow into the positive sector, further decreasing the hiring productivity threshold. Quantitatively, in a calibrated version of the model, the training interval shrinks when an aggregate shock hits, and it enlarges when a sectoral shock hits. This mechanism delivers both the aggregate counter-cyclical impact of training (training incidence going down when a positive aggregate shock hits) and the sectoral fluctuations impact (more training in sectors doing relatively better).

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 provided in Section 3 where we describe the empirical estimation methodology and our findings. Section 4 presents the Mortensen-Pissarides model with training as well as the quantitative analysis from the calibrated model. Section 5 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 of every year.[4] The sample of locations in the frame is stratified by industry (14

[4]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.

strata), region (6 strata), and establishment 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 establishments.

The linked employee component of WES is based on lists of employees made available by the selected establishments. A maximum of twenty-four employees are sampled. In establishments 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 establishments, it can be difficult to infer establishment-specific distributions from the worker side, since only a few workers (sometimes as little as three) are 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 establishment and, more importantly, within the sector − influences the propensity for the worker to get trained.

We define the "extensive margin" of training (based on establishments' self-reported answers) as an indicator of the employer having offered any training to its workers.[5] The intensive margin of training comes either from the percentage of the workforce offered formal classroom training (CT) by each establishment, or from the establishment'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 establishment-specific determinants of training, we follow the literature (*e.g.*, Turcotte, Leonard, and Montmarquette (2003)) by using establishment 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

[5] 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." While we focus our attention on formal classroom training (CT) as a tool for human capital accumulation, we report in the Appendix sensitivity results to using a broader measure of skill improvement, which combines classroom training with on-the-job training (CT+OJT).

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

Variable Description		Mean	Std. Dev.
Classroom Training indicator		0.388	0.487
% Workforce trained		22.3%	34.75
Training expenditures per worker per year (constant 2000 dollars)		\$178.6	\$502.8
Establishment size	Number of workers employed by the establishment	19.28	48.37
Innovation	Adoption of new technology/innovation by the establishment	0.493	0.499
The most dominant sales market of the establishment			
	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

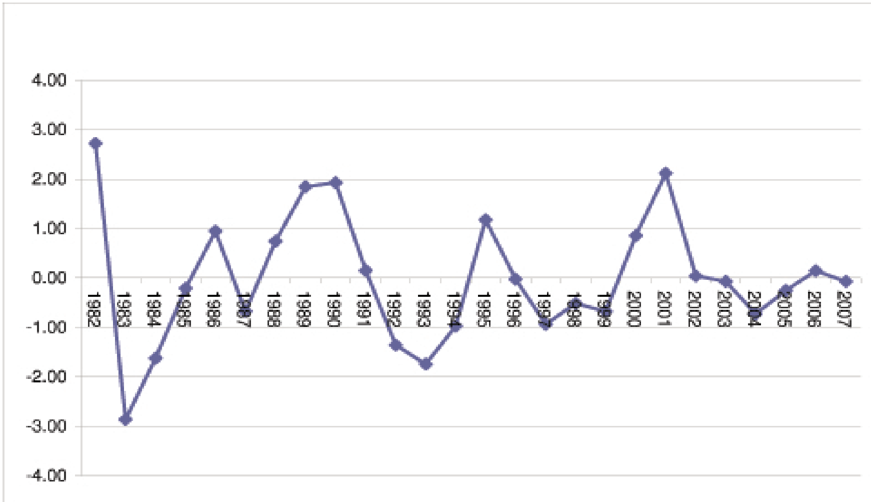
Source: WES 1999-2006, workplace side (2263 establishments).

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 establishments from the agricultural 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.[6]

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

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



3 Empirical Evidence

The extensive and intensive margin analyses are both relevant; the former refers to a model where establishments decide whether to train or not, while the latter refers to a model where selected establishments decide how much to train. When the dependent variable is discrete, we estimate a fixed-effects logit model using the conditional likelihood approach of Chamberlain (1980); when the dependent variable is continuous, we estimate a linear fixed-effects model. In all cases, we allow for unobserved establishment heterogeneity, which can be differenced out:[7]

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

where y_{it} is the training decision of establishment i in period t, α_i denotes establishment-specific heterogeneity, X_{it} are establishment-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:

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

[7] Hausman tests reject random effects in favour of fixed-effects in each specification. Because we do a fixed effect estimation, we cannot use a tobit model to account for the clustering of training observations at zero in the continuous case.

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, workplace side (2263 establishments). Relative size of each sector expressed as the percentage of establishments in the sector relative to total establishments.

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). This may seem like a strong assumption, but, 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 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}. The methodology used for this sensitivity analysis has the added benefit that all potential endogenous regressors, including innovation, are instrumented for. Reassuringly, our story does not change.

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 establishments to train or not to train? The results are in Table 3, where training is defined as an indicator of whether or not the establishment has provided formal classroom training. We provide fixed-effects logit estimates (columns 1 to 3) and fixed-effects OLS estimates (column 4 to 6). 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.[8]

We report standard errors both unclustered and clustered by sectors, to account for the potential correlation in the error terms across establishments within a sector. Insofar as the i.i.d. assumption is violated, clustering the standard errors across sectors provides for consistent inference.[9] Since the robust standard errors are larger, statistical significance can become 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]

The logit coefficient on aggregate output fluctuations (-0.072) is negative and significant, implying that training is counter-cyclical. In other words, establishments 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: establishments 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.

In the OLS estimation, a percentage-point increase in the deviation of log GDP fluctuations from the HP trend will decrease the probability by one percent that an establishment trains. For instance, in our sample the log GDP deviations were 2.12 in 2001 and 0.05 in 2002. Controlling for all other factors, the change in total output relative to its trend would increase training by $2.07 * 0.01$, about two percentage points, all else equal. A one-percentage-point increase in the share of the sector's output in total output will increase the establishment's propensity to train by 0.5 percent. For instance, if the share of business services would increase from 11% (sample average) to 13%, *ceteris paribus* this would lead to an increase in the probability of training of 1% for the business services sector.[11]

[8] Sensitivity results, available from the authors, show that results do no change whether a trend is included or not.

[9] While we only have 10 sectors, which could be a potential source of concern for clustering the standard errors, recent work by Hansen (2007) and Bester, Conley, and Hansen (2009) provide evidence that consistent inference can also be performed when the number of clusters is small.

[10] While it may be interesting to revisit the literature of establishment-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 output deviations, remains unaffected.

[11] While in the fixed-effects OLS model the coefficients represent marginal effects, this is not the case in the fixed-effects logit model, where 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

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

	Fixed-effects logit			Fixed-effects OLS		
	Coef.[a]	Std.Err.	Robust Std.Err.	Coef.	Std.Err.	Robust Std.Err.
	(1)	(2)	(3)	(4)	(5)	(6)
GDP fluctuations[b]	-0.072	0.003	0.036	-0.010	0.003	0.004
Sector to GDP ratio (in %)	0.034	0.002	0.006	0.005	0.002	0.002
Innovation	0.629	0.005	0.142	0.085	0.007	0.019
Market: Canada[c]	0.406	0.011	0.387	0.053	0.014	0.049
Market: World	0.449	0.020	0.598	0.042	0.026	0.071
ln (Firm size)	0.423	0.008	0.132	0.061	0.011	0.023
Multiple locations	0.112	0.006	0.176	0.014	0.008	0.026
Unionized	0.133	0.020	0.364	0.013	0.026	0.037
Staff: % Administrative[c]	0.260	0.020	0.416	0.053	0.026	0.039
Staff: % Managerial	0.560	0.020	0.393	0.085	0.026	0.046
Staff: % Other	1.023	0.021	0.563	0.135	0.028	0.068
Staff: % Sales	0.634	0.021	0.367	0.092	0.027	0.047
Staff: % Production	0.734	0.019	0.661	0.097	0.026	0.087
Staff: % Technical	-0.083	0.018	0.282	-0.029	0.026	0.039
GDP Trend	0.001	0.000	0.001	0.000	0.000	0.000
Constant				-0.073	0.059	0.149

Dependent variable is an indicator of whether the establishment has trained or not. Data from WES 1999-2006, establishment side. The number of observations (establishments) is 8,913 (1,120). Estimation using fixed-effects logit and OLS.
[a]Logit coefficients do not represent marginal effects; OLS coefficients do. The ratio of any two coefficients is the same as the ratio of the two marginal effects (in both specifications).
[b]HP detrended log GDP.
[c]Base category: Local output market; % Professionals in workforce.

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.[12]

cannot average it out across the population (Wooldridge (2005)). Nevertheless, the ratio of logit coefficients equals the ratio of OLS coefficients, which do represent marginal effects (for instance, the ratio of aggregate relative to sectoral impacts is roughly 2 for both logit and OLS coefficients).

[12]In terms of the other establishment-specific factors that influence training, we find that on average more diversified and larger establishments are more likely to train. Establishments that are unionized are also more likely to train, except when adding OJT to the CT definition, when the

Finally, we report in the Appendix two sensitivity checks. In the first one, we explicitly include a proxy for the sector fixed effect, by including on the right-hand side sectoral GDP (from a different data source). These results are very similar to the main ones reported here. Finally, we report sensitivity analysis to defining training slightly differently, by adding on-the-job training to classroom training. Qualitatively, our analysis carries through in this instance as well, though the magnitude of the adjustments is lower.

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 an establishment, and (ii) the training expenditures per worker per year. We start with a specification that includes all establishments in the sample, those who train a positive amount as well as those that do not train (Table 4). We use this specification to benchmark the results from the intensive margin analysis where we select in the sample only establishments that train positive amounts (Table 5). While the magnitude of the coefficients is obviously larger in the positive training specification, the differences are small and do not change the substance of the analysis.[13]

The first column in Table 4 is for the left-hand side variable defined as the percentage of the workforce trained by the establishment. 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 per year (column 4) followed by unclustered and clustered standard errors. A percentage-point increase in the deviation of log 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), one-percentage change in the business cycle measure, GDP fluctuations, decreases the training expenditure by \$4, representing 2.4% of

union coefficient becomes negative (note though that the union incidence in our sample is very low, because 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 the determinants of employer training. See, for instance, Lynch and Black (1998) for the U.S.; Dearden, Reed, and Reenen (2006) for the U.K.; and Turcotte et al. (2003) for Canada.

[13]Moreover, we obtain largely similar results if we restrict the analysis to establishments that train every period.

the annual average training expenditures of \$178.6 (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, or 8.1% (statistically significant). Given sample averages of 22.3% of workforce trained (and \$179 training expenditures per worker), the impacts documented here are relevant, though not huge. Moreover, the average training expenditures are per worker, not per trained worker. It is likely that both the amount spent per trained worker, and our impacts, should be much larger.

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

	% 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 (establishment 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, establishment side. The number of observations (establishments) is 16,208 (2,037). Estimation using fixed-effects OLS.
[a]HP detrended log GDP.
[b]Base category: Local output market; % Professionals in workforce.

The same picture emerges when we focus on establishments 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 (or, 3.9% of the annual average training expeditures). 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 (10.7%) per worker. Establishments which innovate and adopt new technologies train 8 percent more of their workforce and spend \$76 more per worker annually (42.5%) compared to establishments which do not innovate. All coefficients except for the impact of GDP fluctuations on training expenditures are statistically significant.[14]

Overall, the story coming from the continuous definitions of training is similar to the story on training incidence. Establishments 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 (conditional on the same yes/no innovation choice). This is an important finding as the two channels may offset each other. When the economy expands, training on the whole may decrease as the opportunity cost of training has gone up. Nevertheless, in those sectors doing relatively better, overall training may go down by less, or even go up altogether, while in the sectors doing relatively worse, training will go even further down.

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 establishment fixed effects, there remains the possibility that a particular establishment experiences an unanticipated shock specific only to the establishment, and not to the sector or to the economy. While establishments might base their future hiring, innovation, or training decisions on such past idiosyncratic shocks, strict exogeneity rules out this possibility. By relaxing the strict exogeneity assumption and replacing it with the weaker sequential exogeneity below, we allow for covariates, such as innovation or firm size, to depend on past realizations of the idiosyncratic shock:

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

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.

[14] When training is defined as percent of workforce trained, the aggregate and sectoral channels have similar magnitudes (and of course opposite signs); here, the dollar impact is much larger on the sectoral variable than on the aggregate output variable.

Table 5: The Impact of Aggregate and Sectoral Output Fluctuations on Training: Intensive Margin. Sample Restricted to Establishments that 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 (establishment 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, establishment side. The sample is restricted to establishments which train. The number of observations (establishments) is 14,198 (1,784). Estimation using fixed-effects OLS.

[a]HP detrended log GDP.

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

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 (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.[15] Blundell and Bond (1998) show how further additional moment conditions can be obtained by imposing certain restrictions on the initial moment conditions:[16]

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

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 the levels of the regressors as instruments in the difference equation (2), and the lagged values of the 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).[17]

The results from this estimation are in Table 6, where innovation, establishment size, the percentage of workers in each occupation, and the dependent variable, training, are instrumented for. 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. 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

$$u_{it} = \rho u_{it-1} + \varepsilon_{it}.$$

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

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

[17]The Blundell-Bond dynamic system GMM IV estimator is one of the two main approaches used in the literature to help address issues arising from shocks at the establishment level being potentially correlated with future input choices, such as firm size. The other main approach follows Olley and Pakes (1996). In a control-function approach, they model explicitly, in a first-stage structural estimation, the part of the error term u_{it} that is a state variable and impacts the plant's decision rules (and thus may be correlated with X_{it+k}). Also, see Levinsohn and Petrin (2003) for a variation on Olley and Pakes (1996) and a comparison across different estimator choices.

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

	% Workforce Trained				Training Expenditures			
	Contemp.		Lag		Contemp.		Lag	
Training (LHS)			0.21	(0.04)			0.42	(0.15)
GDP fluctuations	-0.85	(1.69)	-0.36	(0.86)	-27.26	(39.26)	-18.20	(23.00)
Sector to GDP ratio (in %)	2.35	(1.16)	-0.92	(0.70)	32.99	(14.92)	10.80	(42.24)
Innovation	9.08	(2.04)	0.04	(1.60)	104.30	(86.05)	-28.66	(61.14)
ln (establishment size)	-3.43	(2.85)	-0.23	(2.84)	-291.18	(657.84)	-93.09	(131.24)
Staff: % Administrative	0.57	(9.00)	-4.13	(8.09)	-384.31	(471.67)	-994.44	(363.89)
Staff: % Managerial	7.63	(8.93)	11.72	(9.46)	154.24	(315.15)	-287.94	(300.75)
Staff: % Other	0.99	(9.30)	7.83	(8.71)	29.33	(301.11)	-582.62	(317.15)
Staff: % Sales	6.97	(10.28)	20.30	(8.83)	337.00	(299.49)	-148.01	(312.99)
Staff: % Production	-0.22	(9.22)	0.33	(7.81)	138.02	(274.34)	-499.83	(316.58)
Staff: % Technical	-6.97	(9.54)	10.02	(8.95)	-46.07	(335.82)	-512.12	(413.94)
Multiple locations	-5.86	(3.63)	2.45	(2.58)	166.18	(57.59)	24.00	(55.23)
Unionized	7.03	(5.88)	1.38	(5.22)	9.00	(73.35)	-91.16	(75.76)
Market: Canada	0.65	(3.31)	-0.91	(2.75)	87.59	(72.28)	-42.90	(96.34)
Market: World	6.30	(5.58)	-3.52	(5.22)	-30.77	(75.50)	-47.12	(88.88)
GDP Trend	-0.22	(0.73)	0.22	(0.72)	-9.21	(25.12)	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	
Number of instruments	107				76			
Overid. test (Hansen)	χ^2(45)=91.10		$P > \chi^2$=0.099		χ^2(45)=61.01		$P > \chi^2$=0.056	

[a]Data from WES 1999-2006, establishment side. 12,642 observations (1818 establishments). Standard errors in parentheses.
[b]HP detrended log GDP.
[c]Base category Local output market; % Professionals in workforce.
Instrumented variables: Training, Innovation, Establishment size, Percent staff in each occupation.

Furthermore, we test the validity of instruments with an IV overidentification test, as by construction we have many instruments, despite adopting a more conservative approach in limiting the size of the instrument matrix.[18] When we report the Hansen test of over-identifying restrictions, which is consistent in the presence of non-spherical errors, the tests show that in both specifications the instruments satisfy the orthogonality conditions implied by the moment conditions. The lag instru-

[18]In choosing the instrument matrix we follow the advice in Roodman (2009a,b). In the standard form, separate instruments are created from each lag of a variable X_t, starting with X_{t-2} (because of the $AR(1)$ assumption), leading to $t-2$ instruments for period t. Summing them up, the total number of instruments is of order $\frac{(T-2)(T-1)}{2}$, a quadratic in T. To reduce the number of instruments we use the "collapse" option with Stata's "xtabond2" command, which collapses the instrument matrix to a single column for each lag distance. In this way the number of instruments is reduced to $T-2$, linear in T.

ments are for the most part not significant, except for the lag dependent variable, which is problematic as it can be indicative of weak instruments; nevertheless, the instruments are jointly significant.

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 establishments in this sector to train. The aggregate channel is still present although less strongly: if the whole economy is experiencing a positive shock, all establishments 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.[19] There is a positive correlation between the sectoral shock and the fraction of new hires who 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.

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 the 1% level.
Data from WES 1999-2006, establishment side.

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

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, establishment size, tenure, and interactions between tenure and education to allow for heterogeneous impacts of tenure across education groups.[20] While we do not know the establishment or sector where a new worker is coming from (nor do we follow workers leaving a establishment), we do know whether a new worker with job tenure of 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 establishment 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 that new entrants into a sector are more likely to get trained.

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).

To illustrate the relationship between output shocks - aggregate and sectoral - and training by establishments, we present below a theoretical model based on the Mortensen-Pissarides matching model (Pissarides (2000)) with heterogeneous workers. The theoretical framework extends the basic Mortensen-Pissarides model to include training decisions by establishments.

[20]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. Note, though, that the coefficients on mostly time-invaryng covariates, such as marital status, are identified only off those individuals who changed status between interviews, and should be interpreted accordingly.

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 establishment	0.057	0.005	1.021	0.006
Medium establishment	-0.016	0.007	1.417	0.008
Large establishment	-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.

4 A Mortensen-Pissarides Model with Training

4.1 The Environment

Consider the following basic matching model in continuous time. Unemployed workers search for a job, and establishments open vacancies whenever they want to fill a job. Keeping a vacancy open implies a cost c. The economy consists of sectors indexed by i (for illustration, we will focus on two sectors only). There is an aggregate pool of unemployed u searching for jobs in all sectors. Establishments can post vacancies v_i in the sector i where they operate. In other words, unemployed workers are fully mobile across sectors while vacancies are not. In each sector, the rate at which unemployed workers and open vacancies meet is regulated by the meeting function $m(v_i,u)$ which depends on the number of unemployed workers and vacancies created in the particular sector i.

Once a worker meets a vacancy, a match-specific productivity α is observed. As in standard Mortensen-Pissarides models, a productive match is formed if and only if α is above the reservation value R_i for that sector. In our extension of the model, training is made available to workers at some cost for the establishment, resulting in increased worker productivity according to a specified human capital function. In equilibrium, two categories of workers will be engaged in productive matches: (i) workers with high ex-ante productivity, who are above a high productivity threshold $\alpha_{\tau,i}$ even without training $(\alpha \geq \alpha_{\tau,i} \geq R_i)$, and (ii) a second group of workers with medium to low ex-ante productivity, who generate enough productive surplus only if they engage in training $(\alpha_{\tau,i} \geq \alpha \geq R_i)$.[21] To keep the model tractable, we assume training to be a continuous activity, so that its cost is a flow cost that has to be paid until the match with training is productive.

A match is terminated by shocks arriving at the rate λ. Training, and the productivity α, are specific to the match: once 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 the same as everybody else in that pool).

Following Pissarides (2000) the meeting function is assumed constant returns to scale $m(v_i,u) = m(1,\frac{u}{v_i})v_i \equiv q(\theta_i)v_i$, where $\theta_i = \frac{v_i}{u}$ is market tightness in sector i. Given the meeting function, the rate at which vacancies are filled can be defined as the probability that the vacancy meets a worker times the probability that a productive match is formed upon meeting,

$$q_i^v = q(\theta_i) \cdot \int_{R_i}^{b} dF(\alpha),$$

where b is the upper limit of the productivity distribution α, and R_i is the productivity reservation value. The rate at which unemployed workers meet and engage in a productive match with a job is respectively given by

$$p_i = \theta_i q(\theta_i) \cdot \int_{R_i}^{b} dF(\alpha).$$

Output is generated as the product between individual productivity and a productivity parameter $y_i = \pi\varepsilon_i$, where π is an aggregate productivity shock for the

[21] Note that there is a third category of workers, those with the lowest productivity (below R_i) who would not meet the reservation threshold even with training. Consequently, they will not get trained and will not engage in a productive match.

overall economy and ε_i is a sector-specific idiosyncratic one. A worker's productivity is α if the worker is not trained and $h(\alpha)$ if the worker is trained. Output is thus given by $y_i\alpha = \pi\varepsilon_i\alpha$ for non-trained workers, and by $y_ih(\alpha) = \pi\varepsilon_ih(\alpha)$ for trained workers. The match is subject to the productivity shock y_i every period; only matches with ongoing productivity above the reservation threshold R_i are preserved.

Wages are set every period by Nash bargaining over the production surplus (net of training costs). Depending on the productivity of the worker and on whether the worker has been trained or not, the wage $w^k(\alpha)$ is paid, with $k = \{0,\tau\}$ superscripting the untrained and trained respectively.

4.1.1 The Value of a Match to an Employer

The value of a job to an employer depends on the productivity specific to that match and on the level of training given to the worker. The value of a match to an employer who does not train $J_i^0(\alpha)$, or who trains $J_i^\tau(\alpha)$, is given by:

$$rJ_i^0(\alpha) = y_i\alpha - w_i^0(\alpha) - \lambda J_i^0(\alpha). \tag{7}$$

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

Here $h(\alpha)$ is a function that describes how productivity increases with training, and it is assumed to be increasing in α, $C(y_i\alpha)$ is the cost of training, and w^τ and w^0 are the wage rates offered to the trained and respectively untrained workers.[22]

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 $\alpha_{\tau,i}$ and it is not required for workers with higher productivity. The asset equation that describes the expected value of a match conditional on $\alpha \geq R_i$ is given by:

$$J_i^e = \int_{R_i}^{\alpha_{\tau,i}} J_i^\tau dG(v) + \int_{\alpha_{\tau,i}}^{b} J_i^0(v)dG(v), \tag{9}$$

where the superscript e indicates the expectation conditional on α being greater than the productivity threshold R_i, and $G(v)$ is the truncated distribution of skill, $G(v) = F(v|v \geq R_i)$.

[22]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.

4.1.2 The Value of a Match to a Worker

The value of a productive match for an untrained or a trained worker is determined by the following asset equations:

$$rW_i^0(\alpha) = w_i^0(\alpha) + \lambda[U - W_i^0(\alpha)], \tag{10}$$

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

Similarly to the previous case for the employer described by equation 9, we can write the expected value of a match for the worker conditional on $\alpha \geq R_i$ as:

$$W_i^e = \int_{R_i}^{\alpha_{\tau,i}} W_i^\tau(v)dG(v) + \int_{\alpha_{\tau,i}}^{b} W_i^0(v)dG(v). \tag{12}$$

4.1.3 The Value of a Vacancy and of Unemployment

The value of setting a vacancy to an employer is

$$rV_i = -cy_i + q_i^v[J_i^e - V].$$

In equilibrium, free entry drives 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 establishments 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

$$rU = z + p_1[W_1^e - U] + p_2[W_2^e - U],$$

where z represents unemployment-contingent income.

4.1.4 Wages and Training

We assume that the wage rates are set following the Nash bargaining rule. Workers with bargaining power β and establishments with bargaining power $1 - \beta$ bargain over the production surplus. After some algebra, and following the same steps as in the standard Mortensen-Pissarides model, we can derive the wages for trained and untrained workers as:

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

$$w_i^0(\alpha) = \beta y_i \alpha + (1-\beta)z + \beta c y_i \theta. \tag{14}$$

The 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,

$$J^0(\alpha_\tau) = J^\tau(\alpha_\tau).$$

Substitution from (7) and (8) we get that $y_i[h(\alpha_{\tau,i}) - \alpha_{\tau,i}] = w_i^{\tau}(\alpha_{\tau,i}) - w_i^0(\alpha_{\tau,i}) + C(y_i \alpha_{\tau,i})$. Finally, substituting from (13) and (14) we get:

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

Note that the reservation training threshold $\alpha_{\tau,i}$ only depends on the cost and benefit of training in sector i, while employment levels or market tightness do not directly affect it.

The Reservation Value for Hiring. The reservation value for hiring is set by the following equation:

$$\max\{J^0(R_i), J^\tau(R_i)\} = 0.$$

Notice that, as long as $R_i < \alpha_{\tau,i}$ (and therefore some training occurs), the relevant condition can be re-written as $J^\tau(R_i) = 0$, or, using (8) and (13), $(1-\beta)[y_i h(R_i) - C(y_i R_i)] = (1-\beta)z + \beta y_i c\theta$.

$$h(R_i) - \frac{C(y_i R_i)}{y_i} = \frac{z}{y_i} + \frac{\beta}{1-\beta}\theta. \tag{16}$$

The reservation productivity threshold R_i depends on the cost and benefit of training, as well as on aggregate conditions in the economy such as labor market tightness, vacancy costs, unemployment contingent income, and workers' bargaining power.

Cost Assumptions. For the model to be consistent with the standard benchmark matching model we need to impose some restrictions. Most importantly, we need to assume that

$$yh(R) - C(yR) - z - \frac{\beta}{1-\beta} yc\theta \tag{17}$$

is, in equilibrium, an increasing function of y, and that

$$yh(\alpha) - C(y\alpha) \tag{18}$$

is also increasing in α. These two conditions imply that, when y increases, the reservation productivity level R decreases as in the benchmark model without training, despite the fact that when y increases, θ also increases.

4.1.5 Training and Unemployment Flows

The rate at which workers enter the unemployment status is given by $\lambda(1-u)$, where u is the total number of unemployed workers in a given period, so that $(1-u)$ is the total number of employed workers. The rate at which workers exit unemployment is given by $\sum_i p_i$. Consider explicitly two sectors only, that is, $i = 1,2$. Then, the rate at which workers get employed is $(p_1 + p_2)u$; that is, workers become employed when they find a match in sectors 1 or 2. Therefore, the rate of change of unemployment is given by: $\dot{u} = \lambda(1-u) - (p_1 + p_2)u$. In a steady-state equilibrium, this change is equal to zero, which leads to an equation for the steady-state unemployment level that is,

$$u = \frac{\lambda}{\lambda + (p_1 + p_2)}. \tag{19}$$

Similarly, we can define the change in the number of workers trained in sector 1, T_1, by $\dot{T}_1 = p(\theta_1)\int_{R_1}^{\alpha_{\tau,1}} dF(v)u - \lambda T_1$, where $p(\theta_1)\int_{R_1}^{\alpha_{\tau,1}} dF(v) = p(\theta_1)[F(\alpha_{\tau,1}) - F(R_1)]$ is the rate at which unemployed workers find a job in sector 1 that requires training, and λT_1 is the exogenous rate of separation of trained workers from sector 1. Imposing the steady state equilibrium, we get:

$$\lambda T_1 = p(\theta_1)[F(\alpha_{\tau,1}) - F(R_1)u. \tag{20}$$

Employment in sector 1 changes according to the flows into employment minus the flows into unemployment: $\dot{E}_1 = p(\theta_1)[1 - F(R_1)]u - \lambda E_1$, which, under steady-state equilibrium, gives:

$$\lambda E_1 = p(\theta_1)[1 - F(R_1)]u. \tag{21}$$

This allows us to write the share of trained workers in sector 1 in steady state as

$$\frac{T_1}{E_1} = \frac{p(\theta_1)[F(\alpha_{\tau,1}) - F(R_1)]u}{p(\theta_1)[1 - F(R_1)]u} = \frac{F(\alpha_{\tau,1}) - F(R_1)}{1 - F(R_1)}. \tag{22}$$

4.1.6 Sectoral Reallocation

To illustrate the adjustment mechanism for the impact of sectoral shocks on training, assume that the productivity in sector 1 increases while in sector 2 it remains constant. That is, y_1 increases while y_2 remains constant. Then, in sector 1,

$$\begin{aligned}
\frac{\partial \frac{T_1}{E_1}}{\partial y_1} &= \frac{[f(\alpha_{\tau,1})\frac{\partial \alpha_{\tau,1}}{\partial y_1} - f(R_1)\frac{\partial R_1}{\partial y_1}][1 - F(R_1)] - [-f(R_1)\frac{\partial R_1}{\partial y_1}][F(\alpha_{\tau,1}) - F(R_1)]}{[1 - F(R_1)]^2} \\
&= \frac{f(\alpha_{\tau,1})\frac{\partial \alpha_{\tau,1}}{\partial y_1}[1 - F(R_1)] - [f(R_1)\frac{\partial R_1}{\partial y_1}][1 - F(\alpha_{\tau,1})]}{[1 - F(R_1)]^2}.
\end{aligned} \tag{23}$$

The sign of the derivative $\frac{\partial T_1/E_1}{\partial y_1}$ is undefined since the two terms in the numerator are both negative. The first term indicates the effect that a change of productivity has on the training threshold $\alpha_{\tau,1}$. Increased productivity implies a lower threshold since the cost raises more than the returns to training; that is, increased productivity raises the opportunity cost of training, making it less profitable for the establishment and thus reducing the share of trained workers in sector 1, $\frac{\partial \alpha_{\tau,1}}{\partial y_1}$.

However, the second term indicates an opposite effect. Increased productivity has a positive effect on employment: sector 1 is now willing to hire more workers because its productivity is higher. This implies that the reservation productivity for hiring decreases with productivity, $\frac{\partial R_1}{\partial y_1}$, and therefore productivity has a positive effect on employment and on training.

The total effect of productivity on training depends on which threshold adjusts downward by more: α_τ or R. If the opportunity cost effect dominates the employment effect, then overall training will decrease; otherwise, it will increase.

In the other sector, which is not directly affected by the idiosyncratic shock, we have:

$$\begin{aligned}\frac{\partial \frac{T_2}{E_2}}{\partial y_1} &= \frac{f(\alpha_{\tau,2})\frac{\partial \alpha_{\tau,2}}{\partial y_1}[1-F(R_2)] - [f(R_2)\frac{\partial R_2}{\partial y_1}][1-F(\alpha_{\tau,2})]}{[1-F(R_2)]^2} \\ &= \frac{-[f(R_2)\frac{\partial R_2}{\partial y_1}][1-F(\alpha_{\tau,2})]}{[1-F(R_2)]^2}. \qquad (24)\end{aligned}$$

To understand the implications of this result, it is sufficient to look at equations (15) and (16). The threshold value of $\alpha_{\tau,2}$ is determined only by the benefit and cost of training and does not depend on any variable that determines the equilibrium in sector 1. Therefore, the threshold value for training in sector 2 does not change because of the change in productivity in sector 1.

The same is not true for the reservation value that determines hiring (R_2), as can be seen from equation (16). An increase in y_1 leads to an increase in tightness θ determined by the improvement in overall productivity.[23] Since the right-hand side of equation (16) increases, the left-hand side of the equation also has to increase, and, if y_2 does not change, this means R_2 has to increase. As R_2 increases when y_1 increases, from equation (24) an increase in the productivity of sector 1, y_1, leads to a decrease in training in sector 2. This negative effect is coming through

[23] This can be seen by looking at the free-entry condition that states that the expected value of a job match has to equate the flow cost of a vacancy divided by the arrival rate of matches to a establishment $q(\theta)$. $q(\theta)$ is decreasing in θ, which implies that θ has to increase when the expected value of a job match increases due, for example, to an increase in productivity.

employment levels: the change in productivity in sector 1 causes a reallocation of workers who flow from sector 2 to sector 1, some of whom will require additional training.

4.1.7 Aggregate Shocks

The easiest way for us to see the effect of aggregate shocks on training is to look at the economy as composed of only one aggregate sector. In this sense, we can re-propose the analysis for sector 1 from above, assuming that is the only sector representing the whole economy. Therefore, we have,

$$\frac{\partial \frac{T}{E}}{\partial y} = \frac{[f(\alpha_\tau)\frac{\partial \alpha_\tau}{\partial y} - f(R)\frac{\partial R}{\partial y}][1-F(R)] - [-f(R)\frac{\partial R}{\partial y}][F(\alpha_\tau)-F(R)]}{[1-F(R)]^2} \tag{25}$$

which simplifyed becomes,

$$\frac{\partial \frac{T}{E}}{\partial y} = \frac{f(\alpha_\tau)\frac{\partial \alpha_\tau}{\partial y}[1-F(R)] - [f(R)\frac{\partial R}{\partial y}][1-F(\alpha_\tau)]}{[1-F(R)]^2}. \tag{26}$$

The final effect of a change in aggregate productivity on training is undefined in aggregate as it was for the one sector affected by the change. Once again, the total effect depends on which effect dominates: the opportunity cost effect, which leads to a decrease in training, or the employment effect, which tends to increase training. However, in the aggregate case we should expect the employment effect to be smaller than in the sectoral case. Equation (16) can once again help us understand this point. Because of the productivity increase, θ also increases. However, y offsets the increase in θ and determines an overall decrease in the reservation productivity level for hiring. The reason why in the aggregate R changes less than in the sectoral case is that θ reacts more to a change in the aggregate productivity.

4.2 Quantitative Analysis

In order to understand better the sectoral and aggregate effects of a change in productivity, we calibrate the model to match certain targets in the data. Then, we use the calibrated model to perform experiments that can quantify the relationship between the productivity shocks and the variables of interest. Note that our model is a steady-state model, and thus the effect of an output shock is given by the comparison between the low-state and high-state output.[24]

[24]In Mortensen-Pissarides type of models the adjustment to the new steady-state is usually fast.

4.2.1 Functional Forms

To keep the model as simple as possible, while still satisfying the conditions in equations (17) and (18), we choose the following parametrization:

$$q(\theta) = \theta^{-0.5} \qquad \text{(that is, } m(v_i,u) = v_i^{0.5}u^{0.5}\text{)} \tag{27}$$
$$\alpha \sim logN(\mu,\sigma^2) \tag{28}$$
$$h(\alpha) = \delta\alpha \tag{29}$$
$$C(y\alpha) = \gamma\alpha^2 y + \eta y^{\psi}. \tag{30}$$

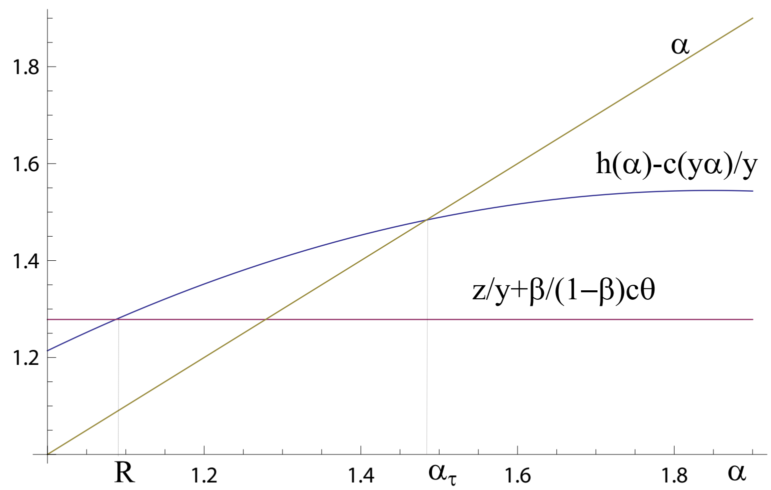
In particular, the cost and benefits of training functions (equations (29) and (30)) are chosen to be consistent with our cost assumptions in (17) and (18) to ensure that the model has a solution for the threshold values of productivity and training. In fact, the cost function needs to be convex in α, while the benefits function needs to be concave in α (in our case, a linear function for the benefits is sufficient and simpler). This will guarantee that the benefits of training are higher than the cost for low levels of ability, and that the net benefit is decreasing in α as required for the thresholds. Given this parameterization, we can re-write equations (15) and (16) as follows (where we omit the sector subscript *i*):

$$\delta\alpha_\tau - \frac{\eta y^{\psi} + \gamma y \alpha_\tau^2}{y} = \alpha_\tau \tag{31}$$
$$\delta R - \frac{\eta y^{\psi} + \gamma y R^2}{y} = \frac{z}{y} + \frac{\beta}{1-\beta}c\theta. \tag{32}$$

Equations (31) and (32) determine the responses in the model to productivity shocks. It is easier to follow the equilibrium adjustments in this model if we summarize (31) and (32) as in Figure 2. While we plot the curves defined by (31) and (32), when discussing the graph it may be easier to refer to the more general equations (15) and (16), which are the threshold equations for α_τ and R before the parametrization has been introduced. In the figure, the 45-degree line represents the right-hand side of the reservation condition for training α_τ in (31) while the flat line represents the right-hand side of the condition for the productivity threshold R in (32). The concave curve is $h(\alpha) - \frac{C(y\alpha)}{y}$, which is the benefit of training net of per-output costs.

When the line for α is above the $h(\alpha) - \frac{C(y\alpha)}{y}$ curve, production takes place without training; therefore, all training must take place to the left of the intersection between the α_τ and $h(\alpha) - \frac{C(y\alpha)}{y}$ lines. The intersection point, α_τ, is the training threshold. Moreover, production with training can only take place when the $h(\alpha) -$

Figure 2: Model plots.



$\frac{C(y\alpha)}{y}$ curve lies above the R line, thus to the right of the intersection point R, which is the productivity threshold. This results in the training interval $[R, \alpha_\tau]$.

An increase in y pushes down the $h(\alpha) - \frac{C(y\alpha)}{y}$ curve, as well as the R line, while leaving the α line unchanged. This results in both intersection points, $[R, \alpha_\tau]$, moving to the left, thereby shifting the training interval to the left. Whether the size of the training interval increases or decreases depends on which threshold decreases by more – R or α_τ. To provide a quantitative answer, we calibrate the model to match the data in certain dimensions, described below, and we perform simulations with the calibrated model.

4.2.2 Calibration

The values of the calibrated parameters are summarized in Table 9. We normalize five parameters to values that are standard in the literature (see the top panel in Table 9). This leaves us with eight parameters (see the bottom panel in Table 9) which we calibrate by matching eight moments from the model with eight moments from the data.

Table 10 provides the moments used in the calibration and reports how well the model moments match the targets in the data. We choose four of the eight targets from the WES data. These targets are: employment tenure, the labor share of value added, the fraction of trained workers, and the training cost as a fraction of value added. We choose three more targets from the Canadian Labour Force Survey, since WES is a survey of workers only. These targets are: unemployment duration,

Table 9: Calibrated Parameters

Parameter	Description	Value
	Normalized	
r	Quarterly Interest Rate	0.01
β	Workers Bargaining Power	0.5
μ	Mean of the Shock Distribution	0
y_1	Structural Productivity 1	1
y_2	Structural Productivity 2	1
	Calibrated	
δ	Training Benefits Parameter	1.0415
γ	Training Cost	0.0246
η	Sensitivity of Training Cost to Productivity	0.0146
ψ	Sensitivity of Training Cost to Productivity Changes	1.160
λ	Job Separation Rate	0.03
c	Vacancy Posting Cost	0.2391
z	Unemployment Benefits	0.4831
σ	St. Dev. of Shock Distrib.	0.4017

the unemployment rate, and the replacement ratio.[25] To calibrate the cost function parameter ψ, we choose the final target such that the model matches the impact that the aggregate channel has on the percentage of workers trained, as documented in the empirical section. Overall, the model matches the data reasonably well on our chosen targets.

4.2.3 Experiments

Table 11 presents the results from the benchmark calibration and from two experiments: (i) when an aggregate shock is introduced, and (ii) when a sectoral shock is introduced. The first two columns of Table 11 show the benchmark case. In the next two columns, we provide the outcome from our first experiment in which we simulate the outcome of a shock that hits both sectors at the same time and with the same intensity. We call this the aggregate shock case. The magnitude of the shock is chosen so that it corresponds to a 1% increase in output. Finally, in the last two columns of Table 11, we simulate the outcome when a positive shock hits Sector 1 and simultaneously a negative shock hits Sector 2. The intensity of the shocks is chosen such that the aggregate average productivity is kept constant to reproduce

[25] We choose a slightly lower replacement ratio to account for the fact that benefits may expire before another job is found, and also that not everyone qualifies for unemployment insurance benefits.

Table 10: Calibration Targets

Data Moment	Data Value	Model Value
Employment Tenure (quarters)[a]	33.3	33.3
Unemployment Duration (quarters)[b]	2	2.19
Unemployment Rate[b]	6.08%	6.17%
Labor Share of Value Added[a]	85%	90%
Share of Trained Workers[a]	22.3%	22.3%
Cost of Training as a Fraction of Value Added[a]	0.76%	0.76%
Unemployment Benefits Replacement Ratio[b]	40%	39%
Change in Training Share from GDP % change[c]	-0.895	-0.891

[a] Source: Authors' calculation from WES 1999-2006. Value-added is computed at establishment level as revenue minus expenditures, where expenditures do not include gross payroll or capital costs.
[b] Source: Canadian Labor Force Statistics.
[c] Impact of aggregate channel from the empirical section (see Table 5.)

a scenario in which there are no fluctuations in aggregate productivity, only in the relative position of the two sectors.

In Experiment 1, when productivity increases on the aggregate, unemployment goes down a little, vacancies increase, θ increases, and both α_τ and R decrease by a little. Because the change in α_τ is slightly larger than the change in R, the training interval shrinks, and the percentage of workers trained goes down, symmetrically in the two sectors.

In Experiment 2, productivity increases in Sector 1 and decreases in Sector 2. Employment goes up in Sector 1 and goes down in Sector 2. Tightness is slightly higher than in the baseline case, as a result of more vacancies being created, despite the fact that unemployment is also slightly higher. In Sector 1, the one hit by a positive shock ε_1, employment increases substantially because workers are drawn in from the bad shock sector in which employment declines. In Sector 1, both reservation thresholds – for productivity and for training – fall compared to the baseline case. Due to the reallocation mechanism, compared to the aggregate shock case, R falls by much more (while the changes in α_τ's are comparable). Subsequently, the percentage of workers trained increases in Sector 1. The opposite holds true for Sector 2 where both R and α_τ increase; since R increases by more, training decreases in Sector 2.

In terms of the impacts documented in the empirical part, we have calibrated the adjustment of the workforce trained to the aggregate shock, but we have left untargeted the impact on training expenditures per worker and the sectoral reallocation channel altogether. We can see how close the predictions from our model are along

Table 11: Experiments

	Benchmark		Experiment 1 Increased overall productivity		Experiment 2 Increased productivity in Sector 1	
	Sector 1	Sector 2	Sector 1	Sector 2	Sector 1	Sector 2
Structural Productivity $y_i = \pi\varepsilon_i$	1.0000	1.0000	1.0104	1.0104	1.0104	0.9850
Percentage of Trained Workers	0.2230	0.2230	0.2141	0.2141	0.2857	0.1282
Employment (in each sector)	0.4692	0.4692	0.4694	0.4694	0.5964	0.3415
Average Productivity	1.2718	1.2718	1.2845	1.2845	1.2771	1.2625
Reservation Productivity R	1.1501	1.1501	1.1500	1.1500	1.1373	1.1671
Reservation Training α_τ	1.1820	1.1820	1.1806	1.1806	1.1806	1.1841
Wages	1.2133	1.2133	1.2255	1.2255	1.2161	1.2074
Training Cost (average per worker)	0.0097	0.0097	0.0094	0.0094	0.0125	0.0055
Training Cost (share of product)	0.0076	0.0076	0.0073	0.0073	0.0098	0.0044
	Aggregate		Aggregate		Aggregate	
Employment Rate	0.9383		0.9388		0.9379	
Unemployment Rate	0.0617		0.0612		0.0621	
Unemployment Duration	2.1915		2.1715		2.2070	
Average Productivity	1.2718		1.2845		1.2718	
Tightness θ	2.7915		2.8410		2.7883	

Note that for the Benchmark case and for Experiment 1, the quantities on the two sectors, Sector 1 and Sector 2, are identical. We report both for consistency, and to ease the comparison with quantities from Experiment 2, which differ across sectors.

these dimensions compared to the empirical impacts documented in Section 3. For training expenditures, the empirical analysis suggests a 3.91% decrease in training expenditures per worker from the aggregate channel, while our model predicts a decline of 3.1% $\left(= \frac{.0094-.0097}{.0094}\right)$ in training expenditures. We find these impacts reassuringly close. Likewise, for the sectoral shock, our model predicts an increase in training expenditures per worker of $29\% = \left(\frac{.0125-.0097}{.0097}\right)$ when the sector is hit by a positive shock, while in the empirical part we document a 10.6% effect of the sectoral shock on training expenditures. While it does seem that the model predicts sectoral expenditure impacts that are too high, the two exercises are not entirely comparable: in the empirical part, the 10.6% impact comes from a one-percentage point change in the sectoral position, while the way we calibrated the idiosyncratic sectoral shock (to keep aggregate productivity constant) resulted in the productivity in Sector 1 changing by 1.9% relative to the benchmark. Once we decrease the magnitude of the training expenditure impacts from the model by half, the model impact gets closer to the empirical impact. Finally, the model over-predicts the impact of the sectoral shock on the fraction of workers trained.[26]

[26] One reason why the sectoral shock generates too much reallocation may be that our sectors are identical to begin with; differences in sector sizes may improve this mechanism.

5 Conclusion

The sectoral analysis is very important when specifying the links between the aggregate business cycle, sectoral idiosyncratic shocks, establishment innovation, and the incidence and intensity of training. We find training to be counter-cyclical: establishments 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. The magnitudes of these adjustments are of similar order.

The experiments performed with the calibrated model have confirmed that we need to take into account both the fluctuations in aggregates and the relative position of each sector in the economy in order to better understand establishments' decisions to train. Taking, for instance, a contraction of the economy, training in the economy as a whole would increase because of the aggregate effect. Nevertheless, some sectors could be contracting much faster than others. Due to the offsetting sectoral effect, training in the fastest declining sectors would increase the least, and could actually decrease.

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 establishments' training decisions over the business cycle, (ii) it quantifies how aggregate and sectoral shocks play into the human capital accumulation decision, and (iii) it could help policy-makers in designing appropriate worker training policies in the future.

From a theoretical standpoint, we highlight the importance for any models of establishment 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 establishments. Moreover, any model used for policy prescriptions regarding training needs to be consistent with both aggregate and sectoral fact; a model consistent with only one or the other may generate misleading predictions.

Appendix A Some Sensitivity Results

Appendix A.1 Adding Sector Proxies

To make sure that our identification strategy does not simply pick up a sector size effect, we perform sensitivity analysis to adding a measure of sectoral fixed effect in our analysis. As sector dummies would get wiped out by the establishment fixed effect, we proxy the sectoral fixed effect by sectoral GDP, as measured in the WES dataset (to construct our main variable of interest, sector to GDP ratio, we used series from the Canadian Statistical Office).[27] Table A1 presents these sensitivity results. The new coefficients differ only very little from those reported in Table 3. As for the sector size, its coefficient is negative (if anything, big sectors tend to train less, as opposed to big establishments, which tend to train more) and not significant once we cluster the standard errors.

Appendix A.2 Definition of Training: Classroom Training Plus On-the-job Training

When the definition of training is expanded to add on-the-job training to the formal classroom training in the training definition, some amount of noise is introduced, because different employers interpret the question of on-the-job training differently (and indeed one could make the case that 100% of the workforce should be counted as receiving on-the-job training). Qualitatively the story carries through. While the magnitude of the two coefficients of interest, aggregate and sectoral output fluctuations, is slightly smaller, their signs and relative sizes still indicate that training moves counter-cyclically with the business cycle and pro-cyclically with the idiosyncratic sectoral shock.

Table A1: The Impact of Aggregate and Sectoral Output Fluctuations on Training Incidence: Extensive Margin. Sensitivity to Adding Sectoral GDP on the RHS.

	Fixed-effects logit			Fixed-effects OLS		
	Coef.	Std.Err.	Robust Std.Err.	Coef.	Std.Err.	Robust Std.Err.
	(1)	(2)	(3)	(4)	(5)	(6)
GDP fluctuations	-0.070	0.003	0.038	-0.009	0.003	0.004
Sector to GDP ratio (in %)	0.044	0.003	0.016	0.006	0.003	0.001
Innovation	0.628	0.005	0.144	0.084	0.007	0.020
Market: Canada	0.404	0.011	0.387	0.053	0.014	0.049
Market: World	0.450	0.020	0.600	0.041	0.026	0.071
ln (Firm size)	0.422	0.008	0.131	0.061	0.011	0.023
Multiple locations	0.105	0.006	0.177	0.013	0.008	0.026
Unionized	0.124	0.020	0.358	0.013	0.026	0.037
Staff: % Administrative	0.259	0.020	0.416	0.052	0.026	0.039
Staff: % Managerial	0.550	0.020	0.391	0.083	0.026	0.046
Staff: % Other	1.017	0.021	0.563	0.134	0.027	0.068
Staff: % Sales	0.635	0.021	0.367	0.092	0.027	0.047
Staff: % Production	0.730	0.019	0.664	0.096	0.025	0.088
Staff: % Technical	-0.088	0.018	0.284	-0.029	0.026	0.039
Sector size	-0.0015	0.0002	0.0017	-0.0002	0.0001	0.0002
GDP Trend	0.001	0.000	0.001	0.000	0.000	0.000
Constant				-0.076	0.061	0.151

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

[27] We thank the editor for suggesting this approach.

Table A2: The Impact of Aggregate and Sectoral Output Fluctuations on Training Incidence. Training is Defined as Classroom Training (CT) plus On-the-job Training (OJT). Extensive Margin.

	Fixed-effects logit			Fixed-effects OLS		
	Coef.	Std.Err.	Robust Std.Err.	Coef.	Std.Err.	Robust Std.Err.
	(1)	(2)	(3)	(4)	(5)	(6)
GDP fluctuations	-0.040	0.003	0.079	-0.005	0.002	0.006
Sector to GDP ratio (in %)	0.022	0.003	0.030	0.003	0.002	0.006
Innovation	0.643	0.005	0.190	0.087	0.007	0.028
Market: Canada	0.580	0.011	0.361	0.071	0.014	0.042
Market: World	-0.491	0.019	0.466	-0.076	0.025	0.063
ln (establishment size)	0.484	0.008	0.203	0.071	0.011	0.032
Multiple locations	-0.132	0.007	0.171	-0.014	0.008	0.019
Unionized	-0.329	0.020	0.433	-0.039	0.025	0.052
Staff: % Administrative	0.154	0.019	0.316	0.029	0.025	0.051
Staff: % Managerial	0.610	0.019	0.424	0.089	0.026	0.058
Staff: % Other	0.265	0.020	0.364	0.038	0.027	0.059
Staff: % Sales	0.669	0.020	0.507	0.080	0.027	0.075
Staff: % Production	0.538	0.019	0.656	0.073	0.025	0.101
Staff: % Technical	0.162	0.019	0.414	0.024	0.025	0.063
GDP Trend	0.0002	0.0000	0.001	0.0000	0.0001	0.0001
Constant				0.352	0.110	0.110

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

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.

Bester, A., T. Conley, and C. Hansen (2009): "Inference with dependent data using cluster covariance estimators," Technical report.

Blundell, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 11–143.

Blundell, R. and S. Bond (2000): "Gmm estimation with persistent panel data: An application to production functions," *Econometric Reviews*, 19, 321–340.

Chamberlain, G. (1980): "Analysis of covariance with qualitative data," *Review of Economic Studies*, 47, 225–238.

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, 397–421.

deJong, D. and B. Ingram (2001): “The cyclical behavior of skill acquisition,” *Review of Economic Dynamics*, 536–561.

Dellas, H. and P. Sakellaris (2003): “On the cyclicality of schooling: Theory and evidence,” *Oxford Economic Papers*, 148–172.

Devereux, P. (2000): “Task assignment over the business cycle,” *Journal of Labor Economics*, 18, 98–124.

Hansen, C. (2007): “Asymptotic properties of a robust variance matrix estimator for panel data when t is large,” *Journal of Econometrics*, 141, 597–620.

King, I. and A. Sweetman (2002): “Procyclical skill retooling and equilibrium search,” *Review of Economic Dynamics*, 5, 704–717.

Levinsohn, J. and A. Petrin (2003): “Estimating production functions using inputs to control for unobservables,” *Review of Economic Studies*, 70, 317–341.

Lynch, L. and S. Black (1998): “Beyond the incidence of training: Employer-provided training,” *Industrial and Labour Relations Review*, 52, 64–81.

Olley, S. and A. Pakes (1996): “The dynamics of productivity in the telecommunications equipment industry,” *Econometrica*, 64, 1263–1297.

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

Roodman, D. (2009a): “How to do xtabond2: An introduction to difference and system gmm in stata,” *Stata Journal*, 9, 86–136.

Roodman, D. (2009b): “A note on the theme of too many instruments,” *Oxford Bulletin of Economics and Statistics*, 71, 135–158.

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*, 39–54.