The Labor Market Impacts of Adult Education and Training in Canada[§]

Shek-wai Hui
Department of Economics
University of Western Ontario
shui@uwo.ca

Jeffrey Smith
Department of Economics
University of Maryland
smith@econ.umd.edu

Final Draft

[§] We thank Human Resources Development Canada for financial support and Statistics Canada for support in data access. We thank Lucie Gilbert for her support, encouragement and patience throughout the completion of this report. We also thank the participants in the HRDC Adult Education and Training Survey Workshop in Ottawa for helpful comments. Any errors are, sadly, our own.

Abstract

In this report, we use the data from the Adult Education and Training Survey (AETS) 1998 to estimate the impact of participating in adult education and training on the employment and earnings of Canadians. We apply methods that assume “selection on observables”, including both standard regression-based methods and propensity score matching methods. We also apply methods based on instruments or exclusion restrictions, including standard instrumental variables estimation and the well-known Heckman bivariate normal selection estimator. This methods aim to deal with “selection on unobservables”. We find that none of the methods we examine produce plausible estimates of the impact of adult education and training, although the methods that assume selection on observables produce more reasonable estimates than those that assume an instrument or exclusion restriction. Based on the results of our analysis, we suggest improvements in the AETS data that would make it a better tool for estimating the labor market impacts of adult education and training.

1. Introduction

Evaluating the impacts of adult training and education is of great value for a number of reasons. To the policymaker, information on the labor market effects of publicly financed adult education and training has obvious implications once placed inside a coherent cost-benefit framework. Similarly, information on the labor market effects of self-financed and employer provided adult education and training provides information on the extent to which existing policies to subsidize or tax such training (or the earnings increments it leads to, if any) lead individual decision makers away from the socially optimal level of participation in these activities. For scholars, information on the impacts of adult education and training provides insight into how individuals and firms accumulate human capital, as well as shedding light on questions of political economy, credit constraints on individuals that may prevent them from making individually and socially optimal investments in human capital, and theories of under-provision of training in the labor market.

The literature distinguishes publicly financed or provided training, especially that for the unemployed or for workers reentering the labor force, from that provided by firms to their employees. There are several reasons for doing so, including the fact that the populations receiving the two types of adult education and training have quite different characteristics, as well as the fact that the content and duration of the training tends to differ substantially. We follow that distinction in our empirical work below and in our brief literature review here.

A large literature exists on the labor market effects of government employment and training programs. Table 25 of Heckman, LaLonde and Smith (1999) lists literally dozens of such studies from numerous countries around the world. In the United States, numerous studies have made use of random assignment methods to produce high quality, credible estimates of the impacts of programs focused on job search assistance, classroom training and wage subsidies. Table 22 of Heckman, LaLonde and Smith (1999) provides a partial list of such studies. The published evaluation record for such programs in Canada is much thinner. The federal and provincial governments commission most Canadian studies for internal use. Few ever see the light of day outside the government and even fewer are ever subject to peer review. Two notable exceptions are Park, et al. (1996) and the widely cited Self-Sufficiency Project, summarized in Michalopoulos, et al. (2000). Riddell (1991), Smith and Sweetman (2001) and Warburton and Warburton (2002) analyze and critique the evaluation of public adult education and training programs in Canada. Fortunately, the types of programs and populations served in the United States are similar enough that their evaluations provide as useful benchmark to which to compare our findings in this study.

The existing literature on the effects of employer-provided training is much thinner. There are several reasons for this. First, governments are, quite reasonably, willing to spend a lot more money evaluating their own programs than evaluating those of private firms. Second, good data on the receipt of employer-provided training is hard to come by. Even when a large survey contains questions relating to employer provided training, there are strong issues of measurement error documented in, e.g., Barron, Berger and Black (1997). Third, while individuals typically participate in government financed training only once, or only at rare intervals when they are unemployed, employer-provided training often continues throughout the lifecycle. Training episodes are often short, on the order of days or weeks, and there are often multiple spells within a year. As documented some below and in Hui and Smith (2002a), these patterns characterize employer provided training in Canada. These features of employer-financed training imply the need for longitudinal rather than cross-sectional data and also make it difficult to know how to code participation – whether in terms of incidence, hours, episodes and so on. The British studies by Arulampalam, Booth and Elias (1997) and by Blundell, Dearden and Meghir (1996) that have attempted to use panel data to study the labor market effects of such training have both wrestled with these issues in depth.

Heckman, Lochner, Smith and Taber (1997) and Carniero, Heckman and Manoli (2002) summarize the evidence from the literature that attempts to evaluate employer-financed adult education and training. One common finding is quite high estimated effects, which are generally attributed to a failure of the available data to completely control for the assumed selection of more able and more motivated employees into training within firms. We did not find any studies along these lines using data from Canada.

In this paper, we estimate the impacts of participation in adult education and training using the data from the 1998 Adult Education and Training Survey (AETS). The AETS is a supplement to the Canadian Labor Force Survey (LFS), and as a result includes all of the information on labor market activity and demographic characteristics included in the LFS. Our analysis has two primary goals. The first is to begin to fill the void in the literature in regard to the labor market effects of adult education and training in Canada. The second is to determine the value of the AETS as a data source for use in evaluation. The second goal is not trivial because the primary focus of the AETS survey instrument is on documenting the types and extent of participation in adult education and training, as well as providing a vehicle for studies of the determinants of participation in such adult education and training, such as Hui and Smith (2002a).

The active literature on econometric methods for evaluating the impacts of treatments such as adult education and training includes a variety of alternative estimation strategies. See Heckman, LaLonde and Smith (1999) and Angrist and Krueger (1999) for recent summaries. The basic problem addressed by all of these estimators is the general absence of data on randomly assignment treatments. In the absence of random assignment, we have observational data, which has the fault that the observed variation in treatment, in our context the observed variation in participation in adult education and training, comes from the (assumed) optimizing choices of individuals. Individuals have information that the analyst does not, and have characteristics that the analyst does not observe. As a result, simple comparisons of the labor market outcomes of participants and non-participants combine the effects of participation with differences due to non-random participation. These differences lead to selection bias.

The literature offers two wide classes of estimators to deal with this problem: those that assume sufficient information in the data to mostly correct for systematic differences between participants and non-participants, and those that assume the absence of such information but instead assume the presence of a variable (an instrument or exclusion restriction) that affects participation but not outcomes in the absence of participation. We utilize two evaluation strategies drawn from each of these broad classes. In the first class, we use standard regression-based methods (the so-called Barnow, Cain and Goldberger (1980) estimator) as well as recently developed propensity score matching methods. In the second class, we use instrumental variables methods as well as the widely known Heckman (1979) bivariate normal selection estimator. In each case, our specifications build on what we learned about the determinants of participation in adult education and training in Hui and Smith (2002a).

As discussed, for example, in Heckman, LaLonde and Smith (1999) and Smith (2000), each of these different estimators makes different assumptions about the processes that generate participation in adult education and training as well as employment and earnings, the two labor market outcomes we examine. At most, the assumptions of one of the estimators we consider matches the data and institutional context we examine here. Our purpose in examining all of them is to allow the data to inform us, in part, regarding which estimation strategies seem most plausible in the AETS context, and also to allow the findings to suggest ways in which the AETS data might be improved for the purposes of impact estimation.

The impact estimates we obtain from all of the econometric methods we apply prove disappointing. The impact estimates for training financed by employers are much too large, while those for training financed by the government are often negative. Theoretical arguments based on expected rates of return, as well as comparisons with alternative estimates in the literature that use better data (in the case of government financed training, often experimental data), cast strong doubt on the estimates obtained here. These poor results hold for all of the estimators we examine, but the results from the instrumental variable and Heckman bivariate normal estimators prove the least credible. Sensitivity analyses indicate that these poor results are robust to modest changes in the specification we estimate, leading us to conclude that the primary problem with the estimates lies in the data rather than in the methods. Put simply, the AETS data lack critical elements necessary to produce credible impact estimates.

The remainder of our study proceeds as follows. Section 2 describes the 1998 AETS data that we use, and defines our measures of training participation and labor market outcomes. Section 3 describes the non-experimental estimation methods we employ to estimate the effects of adult education and training on participants' earnings and employment. Section 4 presents our impact estimates, and indicates why they are problematic in light of the existing theoretical and empirical literature. Section 5 briefly summarizes our suggestions for ways to make the AETS a better tool for estimating impacts; these suggestions are elaborated on in our companion paper Hui and Smith (2002b). We conclude in Section 6 with a summary and some conclusions regarding adult education and training in Canada.

2. Data

2.1 The AETS data

The data we use come from the 1998 Adult Education and Training Survey (AETS) master file. The AETS 1998 is the sixth in a series of similar surveys designed to measure participation in adult education and training, defined as education and training that occurs after the conclusion of formal schooling. The objectives of the survey are to measure participation rates, determine the role of employers in adult education and training participation and provision, and to identify barriers to adult education and training. Statistics Canada collected the AETS data on behalf of Human Resources Development Canada.

The AETS is a supplement to the Labor Force Survey (LFS). The LFS has an overlapping panel design. Each month a new random sample of the LFS population – civilians ages 15 and over – is drawn. Each such sample is called a rotation group. Each rotation group is of roughly equal size, and each one remains in the LFS for six consecutive months, at which point they are no longer followed but instead replaced by a new rotation group. The AETS was administered to five of the six rotation groups in the January 1998 and March 1998 Labor Force Surveys. [1]

The 1998 AETS (hereafter just AETS) consists of five modules, designated A to E.   The questions in Module A collect background information on the respondent. The module also asks whether the respondent received any training or education within the previous year. Respondents who indicate that they did not receive any education or training skip the following three modules, B, C and D, and proceed directly to module E.

The questions in Module B ask about the details of any training or education leading to formal certification of some sort. The AETS calls such education and training "training programs." The questions in Module C ask about the details of any education or training not leading to formal certification, but still intended for career development. The AETS calls such education and training "training courses. The questions in Module D concern courses taken for hobby, recreational, or personal development reasons. We omit the courses reported in Module

D from our analysis due to our focus on training related to labor market outcomes. In each of Modules B, C, and D, the survey collects information on up to five different courses or programs. The information collected on each course or program includes the field of study, the location, the provider, the teaching medium, the duration, whether or not the training was completed, who paid for the training, and what employer support was provided (if any). The survey also collects information on respondents' reasons for taking the training, expectations regarding the training, and opinions of the training's usefulness. All of the questions in Modules B, C and D refer to education and training activities undertaken in 1997.

All respondents complete Module E. This module collects information on labor market outcomes in 1997 for which data are not collected on the LFS. This includes information on the main job in 1997 if it differs from that at the time of LFS completion in 1998. Module E also collects a variety of demographic information including characteristics of the respondents' parents and the respondents' immigration and disability statuses. To supplement the information collected in Module E, the labor force information collected on the LFS is attached to the record of each AETS respondent. In addition, for respondents who report (in Module A) that they did not participate in any education and training in the previous year, Module E includes a series of questions that seek to determine why they did not do so.

The household response rate to LFS is 94.8%, while 85.2% of LFS respondents LFS responded to the AETS. This implies a respectable overall AETS response rate of 80.8%. The 1998 AETS has a total of 33,410 respondents. In order to restrict our attention to those who have

---

[1] The data from March 1998 consists solely of respondents residing in Quebec. This additional

completed their formal schooling, we further restrict the sample to persons 25 to 64 years of age who are not full-time students at the time they complete the LFS. Table 1 shows how these restrictions result in basic analysis samples of 10,748 males and 12,418 females. The samples actually used in some analyses are smaller due to item non-response on particular covariates.

2.2 Defining our measures of training receipt

As noted in the introduction, the literature on adult education and training (hereafter we often just call it "training") distinguishes public and private training for a number of reasons. As, e.g., Hui and Smith (2002a) show for Canada, the populations that receive public and private training differ substantially. Public and private training also tend to differ in their content and in the nature of their providers. Thus, we distinguish between public and private training in this study.

The AETS distinguishes between employer supported training and non-employer supported training. The AETS interprets employer support very broadly, to include such things as unpaid time off for training. In contrast, we feel that someone receiving unpaid time off from his or her employer to participate in a government-sponsored training program should be designated as receiving government training, rather than private training.

Thus, we adopt an alternative definition that (necessarily) relies on the information available in the AETS but instead focuses on who paid for the training. In particular, we define three mutually exclusive categories: employer or union financed training, self-financed training, and government or other financed training. The first category consists of any training paid for, in

---

sample was drawn due to the severe ice storm in Quebec in January 1998.

whole or in part, by an employer or a union. This category dominates the others in the sense that training paid for by both an employer and the government, or by the employer and the individual, is counted only in this category. The second category, self-financed training, includes any training paid for solely by the respondent, along with training provided free of charge. This category may include some training where the government subsidizes (in whole or in part) the training provider, depending on whether or not the respondent recognizes this subsidy in their survey response. It may also include training that receives indirect subsidies in the form of tax credits, transportation assistance, childcare allowances or exemption for job search requirements. The final category is a residual category that includes training the respondent reports as exclusively funded by the government or other sources, such as relatives. The vast majority of the training in the third category is reported as funded exclusively by the government.

As already described, the AETS distinguishes between "programs" and "courses" based on whether or not the training leads, or is intended to lead, to formal certification. Hui and Smith (2002a) show that participation in programs and courses are not mutually exclusive in the AETS data, although the vast majority of participants participate in only one or the other. In addition, they provide some evidence of differences in the determinants of participation for programs and courses. However, given our relatively small sample sizes, and given that the distinction between programs and courses relies on subjective judgements by the respondents, we combine the two types of training in defining our measures of treatment.

Hui and Smith (2002a) show that while some AETS respondents report receiving multiple training spells, whether programs or courses, in 1997, the vast majority of participants report only a single spell. Thus, in our analyses the "treatment" variables consist of indicator (dummy) variables for receipt of employer/union training (courses, programs or both), self-financed training (courses, programs or both) or government/other training (courses, programs or both). Table 2 provides descriptive statistics for the three treatment measures, as well as the underlying distributions in terms of courses and programs for each of the three funding sources.

To get a sense of how much training the treatment represents, Table 3 presents descriptive statistics on the distribution of hours within reported training spells. The top panel indicates the mean duration, as well as the 25th, 50th (median) and 75th percentiles of the distribution for the combined sample of training programs and courses. The information is presented separately for men and women and, within these groups, both overall and separately by the financing source for the training. Four main findings emerge from Table 3. First, employer financed training generally has much shorter durations than government and self-financed training. This holds for both men and women and for both courses and programs. Second, as expected given their definitions, training programs tend to have much longer duration than training courses, although the two distributions do have non-trivial overlap. Third, the mean durations show a remarkable similarity between men and women. Fourth, the data reveal a huge amount of heterogeneity in the intensity of the treatment whose impacts we seek to measure. To take just one example, government financed training for men has a mean duration of 591.5 hours, but the 10th percentile duration is but 12 hours, while the 90th percentile is over 1300 hours.

A number of the training spells in the data remain in progress at the time of the AETS interview in 1998. Persons with a spell of training in progress at that time, which is when our outcomes are measured, are included in the descriptive statistics but omitted from the impact analyses.

2.3 Defining our outcome variables

Studies of the impact of education and training typically focus on their effect on employment and on earnings. We care about employment because the employed generally support themselves, rather than relying on the taxpayer through unemployment insurance or social assistance. Thus, getting people employed represents a goal of many government training and assistance programs. At the same time, for conventional cost benefit analysis, earnings provide a more natural outcome measure. In addition, earnings reflect differences in hours of work and rates of pay between jobs. All else equal, the government (and the trainee!) would prefer that government financed or subsidized training result in full time jobs with higher rates of pay rather than part-time jobs with lower rates of pay.

We follow the literature by defining two outcome measures, one related to employment and one related to earnings. The first outcome measure is employment at the time of the respondent's LFS interview in 1998. The second is the respondent's usual weekly earnings at all jobs as of their LFS interview in 1998. Table 4 reports the means (and, for earnings, the standard deviations) of these variables for the full samples of men and women, and conditional on participation in each type of training (employer/union financed, self-financed, and government/other financed) or receiving no training in 1997.

As discussed in detail in Hui and Smith (2002b), these outcome variables have the very important drawback that they are measured no more than 12 (or 15 in the case of the March 1998 respondents) months after the completion of the training whose effect we seek to measure. In some cases, the lag may be only a month or two, or the training may not yet even be complete. As a result, the outcomes may not fully pick up the earnings and employment effects of training, particularly if it takes some time to find a job following completion of the training. Recent evidence from the California GAIN program presented in Hotz, Imbens and Klerman (2000) suggests that government financed human capital acquisition may have a payoff that does not fully appear for a couple of years after the completion of training.

3. Evaluating the Labor Market Effects of Education and Training

In this section, we lay out a model of labour market outcomes and participation in training. We then describe the assumptions required under different econometric methods of estimating the impact of training on outcomes.

In considering alternative evaluation methods, we are limited by the fact that the AETS is essentially a cross-sectional survey. The AETS collects information on each respondent only once. While the data contain information on total annual earnings for the year 1997, which is the period during which the training it measures takes place, this earnings measure is not comparable to the weekly labor earnings in 1998 measure obtained in the Labor Force Survey. Moreover, for most longitudinal estimation strategies, we want data on the outcome prior to, rather than at the same time as, the training whose impact we seek to estimate. The lack of precise information on the timing of training during 1997 further limits any attempt at using longitudinal methods.

Thus, the data compel us to restrict our analysis to cross-sectional evaluation methods. We consider two pairs of related methods. The first pair of methods relies on the assumption that the data contain information on all of the important factors affecting both labor market outcomes and participation in training. The literature refers to this assumption as "selection on observables." We consider both parametric regression and non-parametric matching estimators that build on this assumption. The second pair of methods allows for selection on unobservables. Both methods require the presence in the data of an instrument (or exclusion restriction). An instrument is a variable that affects participation but not outcomes, other than through its effect on participation. Credible examples of such variables are difficult to come by in this context; we examine the performance of multiple candidate instruments in the empirical work presented in Section 4. As methods are not our primary purpose, our discussion is short and focuses on the main points. Further detail on all of the methods we consider appears in Angrist and Krueger (1999) and Heckman, LaLonde and Smith (1999).

3.1 A model of labor market outcomes and participation in training

The standard human capital earnings function (see, e.g., Becker, 1964, or Mincer, 1974) forms the basis of the outcome models we use to estimate the impact of training on earnings and employment using the AETS. Assuming a linear functional form, we have the outcome equation,

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \ldots + \beta_K X_{Kit} + \delta_{i1} T_{i1t} + \ldots + \delta_{iJ} T_{iJt} + \varepsilon_{it} \,,$$

where Y_{it} denotes the outcome of interest for person "i" in period "t" (earnings or employment in our case), $X_{ikt}, k = 1,..., K$ denote factors such as years of schooling and experience, and $T_{ijt}, j = 1,.., J$ are indicators for receipt of different types of training. Henceforth, given that we have only cross-sectional data, we drop the "t" subscript. For simplicity later on, we define $Y_{1i} = Y_1(X_i, D_i = 1, \varepsilon_i)$ to be the observed outcome with training and $Y_{0i} = Y_i(X_i, D_i = 0, \varepsilon_i)$ to be the observed outcome without training.

Now consider the participation equation.[2] For simplicity, assume for the moment only a single type of training, so that the participation choice consists of taking training or not, and that the training is available only in a single period. Let $Y_i^*(X_i, T_i = 1)$ denote the expected, discounted present value of earnings associated with training. Similarly, let $Y_i^*(X_i, T_i = 0)$ denote the expected, discounted present value of earnings associated with not taking training. Now let $C(W_i)$ denote the expected, discounted present value of the costs associated with taking training, where W_i, which may include elements of X_i, denotes factors that vary the cost of training among persons. Such factors may include age, existing human capital, family characteristics, industry, occupation, job tenure, firm size, region, and so on. Assuming linearity, this gives the training cost function:

$$C_i = \gamma_0 + \gamma_1 W_{1i} + ... + \gamma_L W_{Li} + u_i ,$$

where $W_{1i},...,W_{Li}$ are the individual elements of W_i.

[2] See Hui and Smith (2002a) for more theoretical discussion of the training participation decision,

Risk-neutral individual *i* will take the training if and only if,

$$Y_i^*(X_i, T_i = 1) - C(W_i) > Y_i^*(X_i, T_i = 0).$$

A (very) small amount of algebra allows us to define the expected, discounted gain (or loss) from training as

$$H^*(X_i, W_i) = Y^*(X_i, T_i = 1) - Y^*(X_i, T_i = 0) - C(W_i).$$

We do not observe $H^*(X_i, W_i)$, because we do not observe the counterfactual expected earnings without training for persons who take training or the counterfactual expected earnings with training for persons who do not take training. What we do observe in the data is the decision of whether or not to take training. We can write this decision in the form of a binary choice model, with $D_i = 1$ for persons who take training and $D_i = 0$ for those who do not:

$$D_i = \begin{cases} 1 \text{ if } H^*(X_i, W_i) > 0; \\ 0 \text{ otherwise.} \end{cases}$$

Assuming that $Y^*(X_i, T_i)$ is a linear function of X_i and T_i (as above), that $C(W_i)$ is a linear function of W_i, and that the unobservable components of both have normal distributions centered at zero, yields a reduced form probit model of participation.

The generalization to individuals who are not risk-neutral is straightforward. Simply change the discounted earnings streams above to discounted utility streams. Equally straightforward is the generalization to multiple types of training, so long as we continue to assume that the training takes place in only one period – and assumption consistent with our

---

as well as detailed empirical evidence from the AETS.

cross-sectional data (but not, of course, with reality). In this case, there are multiple possible earnings, or utility, streams, with one for no training and one associated with each available type of training. Each individual chooses the training action associated with the maximum of these discounted values.

Now consider some implications of this simple model of participation and outcomes for the impact estimation undertaken in this model. This is a model of rational training participation. Individuals participate in training when they expect, ex ante, that the benefits will exceed the costs. This feature of the model has several implications. First, it suggests a strong prior belief that the impacts of training, particularly of private training (publicly financed training is sometimes taken for other reasons), will have a positive impact on labor market outcomes. Negative impact estimates will raise suspicion and can be considered an informal specification test of sorts.

The second implication of rational behavior in the context of this model relates to instrument selection. Individuals deciding whether or not to take training are weighing the costs and benefits of doing so. Good instruments will be variables that affect the costs and benefits of taking training without affecting outcomes in the absence of training. Examples of possible instruments suggested by this line of reasoning include variables specifically related to costs, such as distance to the local training center, and variables related to variation in the impacts of training. The third implication of rational behavior relates to heterogeneous impacts. If some individuals gain more (or gain at all) from training and others gain less, we would expect that if individuals can predict their gains to some extent, those we observe taking training will have larger impacts from it than those we do not. This has important implications for policies that seek to increase participation in training, as it suggests care in generalizing estimated impacts of training to populations not presently observed to take it. There may be a reason they are not doing so.

Next, consider the issue of selection bias in the context of the simple model. Some variables affect both participation in training and outcomes in the absence of training. If we fail to take condition appropriately on these variables when estimating the impact of training, our estimates of the impact of training will be biased as the training indicator will proxy for the missing variables that affect both training and outcomes. Two standard examples of such variables are ability and motivation. Both ability and motivation likely have a positive effect on both earnings and participation in private training, which implies a positive bias in the estimated impact of training if they matter and we fail to condition in them. The same is likely true for publicly financed training, although one could make arguments in the other direction.

If we observe the relevant variables that affect both participation in training and outcomes in the absence of training in our data, then we have what Heckman and Robb (1985) refer to as "selection on observables." In this case, including these variables appropriately using the methods discussed in Section 3.3 will suffice to solve the selection problem. If we do not observe the relevant variables, then in terms of the model these unobserved factors result in a correlation between the error terms in the outcome and participation equations, so that $corr(\varepsilon_i, u_i) \neq 0$. In the case, we require methods for "selection on unobservables", which we discuss in Section 3.4. These methods typically require an instrument or an exclusion restriction, which, in terms of our model, is a variable that "belongs" in W_i but not in X_i.

Finally, consider the relationship between this simple, static model and the underlying dynamic process of training participation over the lifecycle. As shown in Becker (1964), it makes sense to take training when young rather than old, as young people have a longer period over which to realize the labor market benefits of their training. This dynamic aspect of the training participation decision can be captured in the static model by including age as a determinant of training.

Another lifecycle issue concerns repeated participation in training. Empirically, we observed individuals taking both publicly financed training (see, e.g., Trott and Baj, 1993, for the U.S.) and private training (see, e.g, Blundell, Dearden and Meghir, 1996 for the U.K.), more than once. Suppose these repeated instances of training are not independent, but instead are positively correlated, perhaps due to unobserved differences in tastes for training. In this case, the impact of current training we estimate may also include the impact of past training episodes we do not observe. To the extent that training has the positive effect that theory suggests it should, this would bias our impact estimates up, if we interpret them strictly as impacts of the training we observe in the AETS.

3.2 Parameters of interest

In a world in which individuals have heterogeneous impacts from training, it is important to consider precisely what the parameter of interest is in evaluating the impact of training.[3] To keep the discussion simple, we again assume for the moment only a single training type, with impact $\delta_i \ (= Y_{1i} - Y_{0i})$ for person "*i*". The extension to multiple training types is straightforward. We consider three possible parameters of interest and briefly discuss the relationships among them.

The average treatment effect is, simultaneously, the effect of training on a randomly selected person in the population of interest or the mean effect on all persons in the population of interest. It is defined as

$$ATE = E(\delta_i).$$

This parameter is of interest in cases where a population will be required (or induced) to participate in training.

The most common treatment effect parameter in the literature is the so-called "treatment on the treated" parameter. This parameter measures the mean impact of training on those observed to receive it in the data. In term of our notation, it is given by

$$TT = E(\delta_i \mid T_i = 1)$$

This parameter is of interest if we want to perform a cost-benefit analysis on training currently being received, whether privately or publicly funded.

---

[3] See Heckman, Smith and Clements (1997) and Heckman and Smith (1998) for extended discussions of heterogeneous treatment effects. For an important early discussion, see Björklund and Moffitt (1987).

The final type of parameter of interest consists of various treatment effects measured at the margin. If we have a binary instrument, the we can define local average treatment effects (LATEs), as in Imbens and Angrist (1994). The LATE is the mean effect on those persons who change participation status when the instrument changes value. It assumes a monotonic response, so that persons do not, for example, become more likely to participate when move farther away from a training center. Each different instrument implies its own LATE, and the LATEs for two different instruments may differ substantially depending on the impacts realized by the persons each instrument induces to participate. If the instrument is a policy variable, such as the tuition for the training, then the LATE may be of great policy policy interest. If we have a continuous instrument, we can define marginal treatment effects (MTEs) as in Heckman and Vytlacil (2001). The marginal treatment effect they define is the effect on the person just indifferent to participating at their current value of the instrument. That is, the marginal person is one for whom $H^*(X_i, W_i) = 0$. Heckman and Vytlacil (2001) show that all of the other common treatment effect parameters can be written as particular integrals of such MTEs.

Finally, we can define other marginal effects not necessarily related to instruments. If we build a new training center in a depressed town, then we can define the impact on the persons who choose to participate in the presence of the new training center who did not participate before when they had to travel to the next town. This treatment effect we refer to as a marginal average treatment effect (MATE). It is not a LATE, because the variation (the new training) center is not an instrument, due to the placement of it in a depressed town where, presumably, outcomes in the absence of training are lower. From the discussion, the policy interest in particular MATEs is clear.

In this paper, we look only at the impact of treatment on the treated for different types of training. We do so for several reasons. First, although it is the parameter most often examined in the literature, estimates for both publicly financed and privately financed training remain somewhat controversial, especially the latter. Second, none of the instruments we examine arise from variation in policy, which is typically necessary for LATEs to be of interest. Finally, as no one is proposing making either public or private training mandatory, potential interest in the ATE parameter in this context is small.

3.3 Cross-Sectional Evaluation Methods Based that Assume Selection on Observables

In this section, we consider two methods based on "selection on observables". That is, both methods assume that we observe in the data all the main factors that affect both participation in training and outcomes in the absence of training.

The most common (and, which is not unrelated, the simplest) method for evaluating the impact of training relies on standard regression methods. For simplicity, we first consider the case of one training type and a common effect of training. The extension to multiple training types is straightforward; we discuss the extension to heterogeneous training impacts below.

Now suppose that X_i includes the standard variables one includes in the human capital model, such as previous schooling and experience (or its proxy, age). But suppose that there remain other factors, not included in the standard human capital model but available in the data, which affect both outcomes in the absence of training and participation in training. Geographic location is a potential example here. These latter variables represent a subset of W_i. We let Z_i denote the union of this subset of W_i with X_i. Under these assumptions, we have that

$$E(\varepsilon_i \mid X_i, T_i) \neq 0,$$

but

$$E(\varepsilon_i \mid Z_i, T_i) = 0. \tag{1}$$

Barnow, Cain and Goldberger (1980) (hereafter "BCG") first derived this motivation for estimating the impact of training using standard regression methods but with a rich set of covariates – rich enough to make the outcome equation error term conditionally mean independent of training.

As discussed in Heckman and Robb (1985) and Heckman and Smith (1996), in a world with heterogeneous impacts, the error term in the outcome equation now implicitly includes the person-specific component of the impact for persons who receive training. That is, the error term includes the difference between the mean impact of treatment on the treated and the individual trainee's impact from training as well as the unobserved component of the outcome in the absence of training. In the BCG set-up, the only major change is in interpretation. The coefficient on the training indicator now just estimates the mean impact of treatment on the treated; under the common effect world it was also an estimate of the average treatment effect.

Like the standard regression estimator, matching assumes selection on observables. However, rather than assuming a functional form for the outcome equation, matching directly compares the outcomes of trained and untrained persons with the same (or similar) values of those variables thought to influence both participation in training and outcomes in the absence of training.

Matching has two advantages relative to the regression estimator just discussed, and one disadvantage. The primary advantage is that it is non-parametric. No functional form assumption from the outcome equation is required to implement the estimator. In standard regression analysis, even if you have the correct covariates, you can still get biased estimates if you assume the incorrect functional form – say by failing to include needed higher order or interaction terms. The second advantage is that you can match on variables that are correlated with the error term in the outcome equation. This is the case because matching only requires that the mean of the error term be the same for trainees and non-trainees with given values of the conditioning variables, not that it be zero. In notation, it requires that

$$E(\varepsilon_i \mid Z_i, D_i = 1) = E(\varepsilon_i \mid Z_i, D_i = 0),$$

but it does not require, as regression does, that both terms equal zero. This is a weaker assumption than assumption (1) above. See Heckman, Ichimura and Todd (1997, 1998) for further discussion. The disadvantage of matching is that, if the linear functional form restriction implicit in regression based analysis in fact holds in the data, then failing to impose it reduces the efficiency of the estimates. Put different, if the outcome equation really is linear, imposing linearity will lead to smaller standard errors on the impact estimates.

The conditional independence assumption (CIA) that justifies matching is given by:

$$Y_{0i} \perp T_i \mid W_i .$$

This assumption implies the balancing condition mentioned in the preceding paragraph. The CIA states that, conditional on W_i, the variables affecting both participation in training and outcomes in the absence of training, participation in training is unrelated to outcomes in the absence of training. Put differently, whatever selection into training takes place, within groups defined by the same values on all the variables in W_i, participation is unrelated to what would happen if the person did not take training. Thus, overall, trainees may have better or worse labor market prospects than non-trainees, but conditional on W_i, their expected labor market outcomes are equivalent. Two technical details deserve note. First, this is the version of the CIA that justifies using matching to estimating the mean impact of treatment on the treated; a stronger version, which also applies to Y_{1i}, is required for estimating the average treatment effect. Second, the variables in W_i may not be factors that may be altered by the trainee in anticipation of taking training (or by a non-trainee in anticipation of not taking training). See Lechner and Miquel (2002) for more on this latter point and Heckman, LaLonde and Smith (1999) for more on the former point.

The problem with matching directly on Z_i is that any set of covariates that plausibly satisfies the CIA is going to be of relatively high dimension. Even if all the elements of Z_i are discrete, the number of distinct combinations becomes large very rapidly, leading to the problem of empty cells – values of Z_i for which we observe participants but no corresponding non-participants to provide the estimated counterfactual. Simply omitting trainees in cells with no non-trainees is not a very satisfying solution. Equally unsatisfying are the various ad hoc cell combination algorithms used in some of the evaluations of the Comprehensive Employment and Training Administration (CETA) program in the U.S. These evaluations are surveyed and referenced in Barnow (1987).

This problem of the potential absence of non-trainees to provide the estimated counterfactual for trainees with certain values of the conditional variables is called the "support problem." The support is a statistical term meaning the set of values for which a density function is non-zero; that is, it is the set of values of a variable that you might observe with positive probability. Along with the CIA, the second main assumption underlying matching is the support condition, given by $\Pr(T_i = 1 | W_i) < 1$ for all possible values of W_i. This condition states that for all values of the conditioning variables, some persons will not participate. Even if this condition holds in the population, it may sometimes fail in finite samples. A more general version of the support condition is required to estimate the average treatment effect; see the discussion in Heckman, Ichimura, Smith and Todd (1998).

Regression-based methods, such as the BCG estimator, implicitly "solve" the support problem through the linear functional form assumption. The functional form assumption fills in where the data are absent. This fact reveals another advantage of matching; it highlights the support condition and makes it clear whether the results obtained were generated by the data, or whether the counterfactuals instead depend heavily on the linearity assumption.

The literature has converged on an alternative solution to the "curse of dimensionality" and the related support problem. Rosenbaum and Rubin (1983) showed that if you can match on Z_i; that is, if Z_i satisfy the CIA, then you can also match on $P(Z_i) = \Pr(D_i = 1 | Z_i)$. This quantity is the probability of participation or "propensity score". This helps solve the problem because $P(Z_i)$ is a scalar – just a real number between zero and one, rather than a vector.

The literature contains a number of different methods of actually implementing propensity score matches. These include nearest neighbor matching (with and without replacement), cell matching, kernel matching and local linear matching. These methods are all consistent in the sense that, as the sample size becomes arbitrarily large, they all give the same answer because in an arbitrarily large sample, all of them rely only on comparisons of trainees and non-trainees with equivalent values of $P(Z_i)$. Detailed discussions of the various methods can be found in Heckman, Ichimura and Todd (1997), Heckman, LaLonde and Smith (1999) and Smith and Todd (2002).

In this paper, where we are concerned with substance rather than methods, we confine ourselves to nearest neighbor matching with replacement, but vary the number of nearest neighbors. Consider first just one nearest neighbor. Nearest neighbor matching without replacement goes through the treated (trainee) observations one by one and, for each one, finds the non-trainee with the nearest (in absolute value) estimated propensity score. That non-trainee becomes the nearest neighbor match for the current trainee and may not be matched to any other trainees. Nearest neighbor matching with replacement proceeds in the same fashion, but allows a given non-trainee to be used as the match for more than one trainee. Matching with replacement reduces the average distance (in propensity scores) between each trainee and his or her matched non-trainee. This should reduce bias. The cost is that if some non-trainees are re-used, this will increase the variance of the resulting estimate. Dehejia and Wahba (1999) clearly illustrate the problem with matching without replacement when the number of comparison non-trainee observations with high probabilities of training is less than the number of trainees with high probabilities of training (as it usually is for obvious reasons). In this case, you get a lot of bad matches. To avoid this, we match with replacement. The formula for the (single) nearest neighbor estimator is given by

$$\frac{1}{n_{\{T_i=1\}}}\sum_{i=1}^{n_{\{T_i=1\}}}(Y_{i1}-\sum_{j=1}^{n_{\{T_i=0\}}}w_{ij}Y_{0j}),$$

where the "i" sum is over trainees, the "j" sum is over the non-trainees, $n_{\{T_i=1\}}$ is the number of trainees, $n_{\{T_i=0\}}$ is the number of non-trainees, and where

$$w_{ij}=\begin{cases}1 \text{ if j = } \arg\min\{|\,\mathrm{P}(W)_i-P(W_j)\,|\};\\ 0 \text{ otherwise.}\end{cases}$$

The generalization to the case of multiple nearest neighbors, each receiving equal weight, is straightforward.

Varying the number of nearest neighbors in the estimation allows us to trade off between the bias and variance in our estimator. Consider switching from using one nearest neighbor to construct the counterfactual for each observation to using two nearest neighbors. The average distance (in terms of propensity scores) between each trainee and the non-trainees used to construct his or her estimated counterfactual mean necessarily increases. At the same time, the number of observations used to construct the counterfactual increases, which reduces the variance of the estimator. The optimal number depends on the density of non-participants. For example, if there are not many more non-participants than participants, there is little gain to using additional neighbors. We experiment with one, two, and five nearest neighbors in our empirical work.

The norm in the economic literature that employs matching is to use bootstrapping methods to estimate the standard errors, for reasons laid out, e.g., in Heckman, Ichimura and Todd (1997,1998). For reasons of time and of computing convenience, we instead report here estimates that do not take account the variance components resulting from the estimation of the propensity score or from the matching itself. Thus, our estimated standard errors are likely downward biased.

3.4 Cross-Sectional Evaluation Methods Based that Assume Selection on Unobservables

We consider two (related) methods that attempt to deal with selection on unobservables. In this setting, we believe that we do not have all of the variables that affect both participation and outcomes in the absence of participation in our data. But, we believe that we have a variable, an instrument or exclusion restriction, which affects participation but does not affect outcomes other than through its effect on participation.

The first of these estimators is the Heckman (1979) bivariate normal selection estimator. This estimator assumes that the error terms in the participation and outcome equations have a joint normal distribution, and that the selection bias results from a non-zero correlation between the two error terms. When the outcome variable is binary, this model corresponds to a bivariate probit model. Technically, the Heckman (1979) model is identified solely based on the joint normality assumption, and no exclusion restriction is required. Extensive experience in the literature in the form of both trial and error and Monte Carlo studies indicates that, in practice, an exclusion restriction is required to ensure the stability of the model. The literature also reveals that the performance of the bivariate normal estimator depends critically on the validity of the normality assumption. Simulation results in Heckman, LaLonde and Smith (1999) show that it also depends on having a strong exclusion restriction – that is, that the variable included in the outcome equation but not in the participation has a substantively important effect on participation. Puhani (2000) summarizes much of the methodological literature on the performance of the bivariate normal estimator.

The bivariate normal model is often referred to as the "Heckman two-step" model, because of the simple two-step procedure to estimate it outlined in Heckman (1979). However, estimation in two-steps is inefficient. As many common software packages (e.g., Stata) now include routines to jointly estimate the participation and outcome equations, it makes sense to do so, and to drop this name for the estimator. There are both common effect and heterogeneous effect versions of the Heckman model; Björklund and Moffitt (1987) first laid out the latter version. For simplicity, we interpret our results in terms of the common effect model.

The instrumental variables (IV) estimator also deals with selection on unobservables. The simplest way to think about how the IV estimator works is in terms of implementing it by doing two stage least squares. In the first stage, the endogenous variable, in our context participation in training, is regressed on the exogenous variables, including both the X_i from the outcome equation and the instrument Z_i. The estimates from this regression are then used to generate predicted values of T_i. Unlike the original values of T_i, these predicted values have a zero correlation with the error term in the outcome equation by construction, because all of the variation is due to variation in the exogenous variables X_i and Z_i. The outcome equation is then estimated using the predicted values of T_i in place of the observed values.

As discussed in Heckman and Robb (1985), the IV estimator requires different and generally weaker assumptions than the bivariate normal estimator. While the difficulty of finding valid instruments has traditionally dominated discussion of this estimator in the literature, the recent methodological literature focuses on two other concerns. The first concern is instrument strength. As the simulations in Heckman, LaLonde and Smith (1999) reveal, weak instruments – in the sense of instruments that have only a substantively weak connection to participation in training – lead to highly unstable estimates. Bound, Jaeger, and Baker (1995), and the literature that has followed from it, also show that IV estimates based on weak instruments can be biased. As a result of this line of research, empirical work typically now pays more attention to the importance of the instruments in predicting the endogenous variable.

The second concern is with the interpretation of the IV estimates in a world with heterogeneous treatment effects. An instrument may valid in the sense that it is not correlated with the outcome equation unobservable ε_i, but may nonetheless be correlated with the person-specific component of the impact. In this situation, the interpretation of the estimate produced by instrumental variables is problematic. Under certain assumptions, the resulting estimate may be a LATE rather than an estimate of the impact of treatment on the treated, which is generally the parameter of interest. Heckman (1997) discusses this issue in detail.

3.5 Assumptions and Data

Table 5 summarizes the variables available in our data set. How well do these variables satisfy the requirements of the estimators discussed in Sections 3.3 and 3.4? Can we plausibly employ the BCG, matching, IV and bivariate normal estimators using our data, given what we know from the literature about the determinants of participation in training and of outcomes in the absence of training? This is the question addressed in this section.

For clarity, Table 6 divides the variables in Table 4 into three sets, A, B and C. Set A represents variables we think influence both outcomes and participation, and which are plausibly exogenous (not correlated with the error term) in the outcome equation. This set includes the standard variables included in human capital earnings functions. In such models, education captures differences in human capital across individuals due to differences in the amount of formal schooling obtained, while experience captures the effects of human capital obtained through learning by doing. In the absence of a measure of years of experience, as in the AETS, age serves as a proxy measure.

In practice, other variables affecting labor market outcomes also appear in human capital earnings functions (and employment equations). Such variables include regional dummies that capture geographic differences in labor market conditions (and price levels, in the case of earnings), and an immigrant status indicator to capture differences due to language skill, educational quality or discrimination. We follow the standard practice of estimating the impacts of training separately for men and women, rather than just including an indicator variable for sex. Familiar differences in lifecycle labor supply due to childbearing and other factors, and the resulting frequent statistical rejection of models that pool men and women, motivate this practice. Thus, there is no need to include a sex variable in set A. We include the variables in set A in the outcome equations of all the regression-based models we estimate, and include them in all specifications of the propensity score for the matching estimates.

Set B represents variables that we expect to affect both participation in training and outcomes in the absence of training, but which we think are endogenous (correlated with the error term) in the outcome equation. These variables include mainly characteristics of the 1997 job, including firm size categories, and indicators for union membership and for professional and blue-collar jobs. As some sample members were not working in 1997, we include indicator variables for missing values on the job characteristics, along with indicators for any employment and full-time employment. Due to our concern with endogeneity, we do not include these variables in any of the regression-based models, as including them would likely bias the estimates we obtained. However, because we do not think that anyone would deliberately alter any of these variables in anticipation of taking training, we feel confident in including them in some of our propensity score matching specifications.

Set C represents variable we think can plausibly be used as instruments for training participation. These variables include various measures related to the respondent's spouse, if one is present, and his or her children. The most problematic variable here is whether or not there is a spouse present, as there is a (controversial) literature that suggests a married wage premium for men. If such a premium exists, marriage does not constitute a good instrument. Waite and Gallagher (2001) summarize the literature on this. Because of this concern, we repeated our analysis deleting the marital status variables from Set C, and obtained qualitatively similar results. We use the variables in Set C as instruments for our IV estimates and as exclusion restrictions for our bivariate normal estimates.

4. Estimates of the Labor Market Impact of Training

In this section we present our estimates of the labor market impacts of training using the data from the AETS. For ease of presentation, we divide our presentation based on the two outcome variables of interest – employment and earnings. A short summary of the impact estimates ends the section.

We examine several alternative models because we do not have strong a priori beliefs about how well the assumptions for any of the models correspond to the population particpation and outcome processes for adult education and training in Canada, given the data we have available from the AETS. As discussed in, e.g., Smith (2000) there is no one econometric evaluation estimator that is always the correct one, regardless of the available data and the substantive context under consideration, although one could not be faulted for drawing this conclusion from much of the literature. As such, we consider estimates from several alternative estimators here.

4.1 Impacts of Training on Employment

Table 7 lists our estimates of the impact of any training on employment as of the 1998 LFS survey. For ease of presentation, estimates of the other parameters in each model appear in the appendix, which also presents estimates of the propensity score models, of the participation equation in the bivariate probit and bivariate normal models, and the first stage estimates from the IV estimation.

Each row of Table7 corresponds to a different estimator, with the estimators listed in the order in which we discussed them in Section 3. All of the estimates should be interpreted as estimates of the impact of treatment on the treated, as defined in Section 3.2. The first row presents estimates from a linear probability model (LPM) version of the BCG estimator, using the variables in Set A variables as covariates. The second row presents the same model but estimated as a probit. The impact estimate is the marginal effect evaluated at the mean of the covariates, and so is comparable to the estimate from the LPM and the other models.

The next six rows present various propensity score matching estimates. There are three pairs of matching estimates. Each pair uses the same number of nearest neighbors but varies the set of conditioning variables. We consider estimates based on one, two and five nearest neighbors. Within each pair, we present estimates using scores that include only the variables in set A and scores that include the variables in sets A and B. Comparisons within each pair show the effect, if any, of conditioning on a richer set of covariates – covariates that could not be included in the regression-based models due to their correlation with the outcome equation error term.

The final two rows present estimates from models that attempt to take account of selection on unobservable variables, such as ability or motivation. The penultimate row in the table presents two-stage least squares estimates obtained using the variables in set C as instruments for training in a regression of employment on the variables in set A plus the training indicator. The estimated standard errors for these estimates take account of the additional variation induced by the estimation of the first stage; a full set of estimates from the first stage (including the first stage F-statistic) appear in the appendix. The final row of Table 7 presents estimates from a bivariate probit model, using the variables in set C as exclusion restrictions. The impact estimate from the bivariate probit model is a marginal effect estimated at the mean of the covariates in the outcome equation.

Tables 8A and 8B present a similar set of estimates for each of the three types of training: employer-financed, self-financed and government-financed. The estimators here are similar to those in Table 7, but with a few changes. To estimate the effects of the three types of training in the BCG model, we simply include indicators for the three trainings in the LPM and probit models in place of the single indicator for any training used in Table 7. For the matching estimates, we generate separate impact estimates for each training type using the trainees with that type of training as the treatment group and those sample members with no training of any type as the comparison group. We follow the same strategy for the IV estimates as we do for the BCG estimates, by replacing the single training indicator with separate indicators for the three training types. We can do this because set C contains more than three variables; as is well known, in order to satisfy the rank condition, at least one instrument is required for each endogenous variable. Finally, we do not present any estimates by training type for the bivariate normal model; the generalization of this model to multiple training types would require the joint estimation of a probit outcome equation and a multinomial probit model with four alternatives, a numerical task which is beyond the scope of this report.

Before turning to the estimates themselves, we briefly consider the support condition for the matching estimates. Figures 1A to 1H show histograms of the distributions of estimated propensity scores for participants and non-participants in any training, and for employer, government and self-financed training. Consistent with our estimation procedure, we present separate figures for men and women. Support problems arise when there are intervals of the propensity score that have positive density for participants but zero or very low density for the non-participants. A support problem means is that there are no (or at least not very many) non-participants to utilize in constructing the estimated counterfactual for the participants. The figures show no support issues for the matching estimates of any training and of employer financed training for both men and women. Minor support problems arise at some higher values of the probability of participation for government financed training and self-financed training. Because the fractions of participants in these regions are small, the support issue is likely of second-order importance relative to the other concerns we discuss in what follows.

The estimated impacts of any training presented in Table 7 reveal several important patterns. First, the estimates obtained from the linear probability model (LPM), the probit (translated into a marginal effect for comparability) and matching on the variables in set A (the same variables included in the LPM and the probit) all give about the same answer. This indicates that the linear functional form restriction implicit in the LPM and the probit is not crucial here. Second, matching on the variables in set B, the variables related to employment status and job characteristics in 1997, reduces the impact estimates by well over 50 percent. This illustrates the value of being able to condition on "endogenous" variables in matching. The variables in set B clearly capture differences between trainees and non-trainees not captured by the variables in set A.

Third, the impact estimates from the IV and bivariate probit are huge – far too huge to be plausible given the magnitude of the treatment being evaluated, which has a mean of around 320 hours according to Table 3. We consider the plausibility of all of the estimates in more detail in our discussion of the earnings impacts below. The bizarre values of the IV and bivariate normal estimates indicate that set C, which contains the most promising candidates for valid instruments in the AETS data from a theoretical standpoint does not in fact contain valid instruments. To verify this, we repeated the analysis using each variable in set C as a single instrument in turn; doing so did not change the qualitative findings. We also performed "Hausman" tests of the instruments. The basic procedure here is to assume that all but one of the candidate instruments are valid and test the remaining one conditional on the others. This procedure is then repeated for each candiate instrument (or set of instruments in the case of dummies for categories of a common variable). Surprisingly, we could not reject the null of valid instruments in every case but one. The weakness of this test of course, is that it needs at least some of the instruments to be valid ones to work, and it is not clear given our estimates that we meet this requirement.

Fourth, and finally, the estimates for men and women are surprisingly similar.

Turning now to the estimates by type of financing, the picture becomes even more interesting. The most obvious finding in Tables 8A and 8B is that the estimates for any training in Table 7 disguise a substantial amount of heterogeneity by type in terms of the impact of training. In Tables 8A and 8B, employer training has a positive impact, which is quite large with most of the estimators, government financed training has modest to large negative impacts, and the impact of self-financed training varies in sign depending on the estimator in question for men while having small but consistently postitive impacts for women.

Most of the other patterns observed in Table 7 carry over to Tables 8A and 8B. The matching estimates using the set A variables remain very similar to the estimates from the linear probability models and the probit models using the same covariates. Adding the set B variables again reduces the magnitude of the impact estimates – even more substantially here than in Table 7. For example, For men, the impact estimates for employer traning fall from around 12 percentage points to around one percentage point when matching on variable sets A and B rather than just variable set A. Once again, the theoretical advantage of matching, which allows comparisons conditional on variables that would be endogenous in a regression framework, has important practical consequences. As in Table 7, we obtain implausible estimates using the IV estimator, which we again interpret as reflecting poorly on the available instruments in the AETS. The general patterns in the estimates remain suprisingly similar for men and women.

In terms of substance, only the matching estimates using variable sets A and B appear reasonable on the face of it. The others are all too large in magnitude to be plausible given the average duration of the training whose impact we seek to measure. Even with matching on the set A and B variables, we still obtain negative impact estimates for government financed training. While there is some experimental precedent for negative impacts from government training – as in the findings for male youth in the U.S. National Job Training Partnership Act Study (NJS) documented by Bloom et al. (1993), we hesitate to jump to the same conclusion for Canadian government financed training here, especially since positive impacts were found in the NJS for adults, which correspond more closely to the age group under consideration here.

4.2 Usual Weekly Earnings

Tables 9 for any training and 10A and 10B for training by type present estimated impacts on usual weekly earnings. These tables repeat the basic structure of Tables 7, 8A and 8B in terms of the estimators presented with only a couple of minor differences reflecting the change from a binary outcome variable to a continuous one. First, instead of reporting a LPM and a probit for the BCG model, we report a single linear regression. Second, the bivariate normal estimator now combines a probit participation equation with a linear outcome equation; the estimates presented are from full information (joint) estimation of this model, rather than from the traditional two-stage approach due to Heckman (1979). Given the availability of the more efficient (i.e., smaller standard errors) full information version of the model in widely used software packages, there is no reason to use the two-step method. As in the preceding section, we do not report estimates of the bivariate normal model for the individual training types; the extension of the normal model to the multiple treatment case as in Lee (1983) or Dubin and McFadden (1984) is beyond the scope of this report.

Not surprisingly, the earnings impacts in Table 9 for any training show much the same patterns observed already for the employment impacts. This includes the similarity of the linear regression estimates and the matching estimates using the set A variables, the steep decline in the matching estimates when conditioning on the set A and set B variables compared to just the set A variables, the extreme implausibility of the IV and bivariate normal estimates, and the general similarity of the estimates for men and women. These same patterns recur again when we break the training down by type of financing in Tables 10A and 10B. We also again observe strong differences by financing type, with large positive estimated earnings impacts for employer training along with generally large negative estimates for government financed training. The one difference to the empolyment impacts is that earnings impacts for self-financed training are typically nearly as negative as those for government-financed training.

4.3 Plausibility of the Impact Estimates

In the context of earnings as a dependent variable, it is somewhat easier to consider the plausibility of the estimated impacts. Consider two benchmarks. The first is the literature on the earnings impacts of a year of formal schooling. Card (1999) provides a recent summary of this literature. A typical value for the earnings impact of a year of formal schooling from a log wage equation with some attempt to control for ability bias is around eight percent. Now consider the mean earnings levels of the individuals in the AETS data. As shown in Table 4, the overall mean for men is about $500 per week while the average for women is about $300 per week. The average duration of a year of formal schooling is about 32 weeks; the average duration of training in the AETS as shown in Table 3 is about eight weeks (= 320 hours / 40 hours per week). Thus, if the training measured in the AETS is about as effective as formal schooling per hour, it should have an earnings impact of about two percent (= 8/32 * 0.08). For men, this corresponds to about $10 per week and for women it corresponds to about $6 per week.

The second benchmark comes from a rate of return calculation. The primary cost of training is likely to be its opportunity cost. For men, eight weeks of training has a rough opportunity cost of $4000 (= $500 per week * eight weeks), while for women it has a rough opportunity cost of $2400 (= $300 per week * eight weeks). Assume a five percent annual rate of return and that the earnings impacts of training persist indefinitely. Then a $4000 investment should yield an impact of about $200 per year or about $4 per week. Similarly, a $2400 investment should yield an impact of about $120 per year or about $2.50 per week.

Obviously, both these benchmarks are fairly crude. As the Card (1999) chapter shows, there remains substantial disagreement about the earnings impact of a year of formal schooling. The crude rate of return calculation ignores differences in opportunity costs for persons taking different types of training (these are clear from the differences in average weekly earnings in Table 4), as well as the fact that individuals will tend to take training when opportunity costs are transitorily low, whether due to slack demand at a firm or due to unemployment at the individual level. Taking these caveats into account, however, both benchmarks still strongly suggest the implausibility of estimates exceeding $10 per week.

The most plausible estimates in Tables 9, 10A and 10B are those from matching on both variable sets A and B. A comparison of these estimates to the benchmarks, however, reveals that even their magnitudes are, with a couple of exceptions, far too large to be plausible. For example, in Table 9, these estimates range from $33 to $45 for men and from $20 to $28 for women, depending on the number of neighbors utilized in the nearest neighbor matching. This is five to ten times larger than our benchmarks suggest as reasonable estimates.

4.4 Summary of the Impact Estimates

We can summarize our findings along two lines: variation in the estimates across alternative econometric evaluation estimators and the substantive meaning and plausibility of the estimates. In regard to the former, our findings suggest that we lack valid instruments in the AETS, with the result that we obtain wildly implausible estimates using the IV and bivariate normal estimators. Conditional on covariate set A, we find little difference between standard linear estimators and non-parametric matching estimates. The gain from matching comes when we match on variable set B, which could not be included in the linear models due to endogeneity concerns. Including the set B variables substantially reduces the magnitude of the estimates in all cases.

In terms of substance, only the matching estimates using both the set A and set B variables come close to being plausible when considered in light of the evidence from the literature on the earnings impact of a year of formal schooling, or when considered in terms of standard rates of return on investment. Even there, they remain many times larger than these external benchmarks suggest they should be. The exception is the case of government-financed training, where the point estimates are negative and statistically significant in most cases. Although the evaluation record for public employment and training programs is not a pretty one –

see, e.g., the survey in Heckman, LaLonde and Smith (1999) – these results are too negative even relative to the low standard set by the literature.

Thus, our overarching conclusion from the empirical work is that the AETS requires some modification for use as a tool to estimate impacts, in addition to its primary role as a tool for measuring the incidence and nature of adult education and training. In a companion paper, Hui and Smith (2002b), we discuss in detail proposed changes in the survey to make it into a more effective tool for evaluation. Without those changes, our findings suggest extreme caution in the generation and interpretation of impact estimates from these data. In the following section, we briefly summarize the primary themes in Hui and Smith (2002b).

5. Problems with the AETS for Impact Evaluation

Hui and Smith (2002b) provide a detailed catalogue of suggestions for improving the usefulness of the AETS for the task of estimating the labor market impacts of adult education and training in Canada. In this section, we summarize their four primary themes.

The first theme in Hui and Smith (2002b) is a general one. Better information on timing would substantially improve the AETS. At present, it is impossible to when a spell of training ends in relation to the survey date at which outcomes are measured. With the present survey structure, this difference could range from one month to 14 months. This makes a difference, both in terms of interpreting the obtained impact measures and for attempts to measure the time path of impacts. A related issue concerns the timing of job spells in the year prior to the survey.

Better timing information here would allow the researcher to always link spells of training to particular jobs.

The remaining three themes all center on providing data for particular econometric evaluation methods. Thus, the second theme in Hui and Smith (2002b) concerns the value of obtaining information on additional covariates for use with estimators that assume selection on observables. Such variables could include measures of family background (e.g., parental education), more detailed measures of past degree and certificate attainment, or a measure of ability. The third theme concerns the value of obtaining repeated outcome measures (or at least one outcome measure) prior to training, in order to allow the utilization of longitudinal evaluation methods. Finally, the fourth them concerns the value of having more credible instruments associated with the AETS data. These could include information from respondents on the nearest training center or on subsidies available at their firm for training, or it could include matched information on external sources on variables such as tuition at public colleges in each province. The potential value of each of these lines of improvement in the AETS is clear from the analysis presented in this paper.

6. Conclusions

Our findings lead to a number of important conclusions. First, the impact estimates from the AETS exhibit extreme sensitivity to the choice of non-experimental impact estimator. Careful consideration of the plausibility of alternative econometric evaluation methods is crucial in using these data to obtain estimated impacts of adult education and training.

Second, our findings suggest the absence of a plausible instrument for training among the variables available in the 1998 AETS. As a result, we obtain extremely large impact estimates when applying the IV and bivariate normal estimator – estimates far too large in magnitude to be believed.

Third, we find that there is little difference in estimates obtained from non-parametric matching methods and traditional linear models such as regression or probit when estimated using a common set of covariates. The advantage to matching in this application is that it allows conditioning on variables related to employment status and job characteristics in the previous year. Including such variables in a linear model would raise endogeneity concerns. The inclusion of these variables – denoted the set B variables in our analysis – substantially reduces the magnitude of the estimated impacts. Indeed, the only really plausible estimates we produce come from matching on the combined A and B covariate sets. These findings highlight two methodological points already present in the literature but often times ignored: first, what you match on matters a lot for the answer you get, and matching may produce very poor estimates when applied with a covariate set for which the conditional independence assumption that justifies matching fails to hold.

Fourth, even the most plausible of the estimates, those obtained by matching on both variable sets A and B, where the latter includes variables measuring employment status and job characteristics in the previous year, lack plausibility. When compared to benchmarks based on estimates of the return to a year of formal schooling or upon simple rate-of-return calculations, these estimates appear at least five to ten times too large.

Finally, the overall conclusion from our study is that the AETS as presently constituted is not a very good vehicle to use in evaluating the impact of adult education and training. This is perhaps not surprising, as it was designed to measure the incidence and variation in type of adult education and training. However, it is a fact worth knowing. In our view, however, with modifications along the lines suggested in Hui and Smith (2002), it could become such a vehicle. Having a data set, whether the AETS or some other, that could be readily used to evaluate the impacts of adult education and training would be an asset to both policy and research in Canada.

References

Angrist, Joshua and Alan Krueger. 1999. "Empirical Strategies in Labor Economics." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Volume 3A. Amsterdam: North-Holland. 1277-1366.

Arulampalam, Wiji, Alison Booth and Peter Elias. 1997. "Work-Related Training and Earnings Growth for Young Men in Britain." *Research in Labor Economics*. 16: 119-147.

Barnow, Burt . 1987. "The Impact of CETA Programs on Earnings: A Review of the Literature," *Journal of Human Resources*. 22: 157-193.

Barnow, Burt, Glen Cain and Arthur Goldberger. 1980. "Issues in the Analysis of Selectivity Bias." In Ernst Stromsdorfer and George Farkas, eds., *Evaluation Studies Review Annual*, Volume 5. Beverly Hills, CA: Sage Publications. 43-59.

Barron, John, Mark Berger and Dan Black. 1997. "How Well Do We Measure Training?" *Journal of Labor Economics*. 15(3, Part 1): 507-528.

Becker, Gary. 1964. *Human Capital*. New York: Columbia University Press.

Björklund, Anders and Robert Moffitt. 1987. "Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Review of Economics and Statistics*. 69(1): 42-49.

Bloom, Howard, Larry Orr, George Cave, Steve Bell and Fred Doolittle. 1993. *The National JTPA Study: Title II-A Impacts on Employment and Earnings at 18 Months*. Bethesda, MD: Abt Associates.

Bound, John, David Jaeger, Regina Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variables is Weak." *Journal of the American Statistical Association*. 90: 443-450.

Blundell, Richard, Lorraine Dearden and Costas Meghir. 1996. *The Determinants and Effects of Work-Related Training in Britain*. London: Institute for Fiscal Studies.

Card, David. 1999. "The Causal Effect of Education on Earnings." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Volume 3A. Amsterdam: North-Holland. 1801-1864.

Carneiro, Pedro, James Heckman, and Dayanand Manoli. 2002. "Human Capital Policy." Unpublished manuscript, University of Chicago.

Dehejia, Rajeev, and Sadek Wahba. 1999. “Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs.” *Journal of the American Statistical Association*. 94(448): 1053-1062.

Dubin, Jeffrey and Daniel McFadden. 1984. “An Econometric Analysis of Residential Electrical Appliance Holding and Consumption.” *Econometrica*. 52(2): 345-362.

Heckman, James. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*. 47(1): 153-161.

Heckman, James. 1997. “Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely Used Estimator.” *Journal of Human Resources*. 32(3): 441-461.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith and Petra Todd. 1998. “Characterizing Selection Bias Using Experimental Data.” *Econometrica*. 66(5): 1017-1098.

Heckman, James, Hidehiko Ichimura and Petra Todd. 1997. “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program.” *Review of Economic Studies*. 64(4): 605-654.

Heckman, James, Hidehiko Ichimura and Petra Todd. 1998. “Matching as an Econometric Evaluation Estimator.” *Review of Economic Studies*. 65(2): 261-294.

Heckman, James, Robert LaLonde and Jeffrey Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Volume 3A. Amsterdam: North-Holland. 1866-2097.

Heckman, James, Lance Lochner, Jeffrey Smith and Christopher Taber. 1997. “The Effects of Government Policy on Human Capital Investment and Wage Inequality.” *Chicago Policy Review*. 1(2):1-40.

Heckman, James and Richard Robb. 1985. “Alternative Methods for Evaluating the Impact of Interventions.” In James Heckman and Burton Singer, eds., *Longitudinal Analysis of Labor Market Data*. New York: Cambridge University Press for Econometric Society Monograph Series. 156-246.

Heckman, James, and Jeffrey Smith. 1996. “Experimental and Nonexperimental Evaluation.” In Günther Schmid, Jacqueline O’Reilly, and Klaus Schömann, eds., *International Handbook of Labour Market Policy and Evaluation*. Brookfield, VT: Edward Elgar. 37-88.

Heckman, James, and Jeffrey Smith. 1998. "Evaluating the Welfare State." In Steiner Strom, ed., *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial*. New York: Cambridge University Press for Econometric Society Monograph Series. 241-318.

Heckman, James, Jeffrey Smith, and Nancy Clements. 1997. "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts." *Review of Economic Studies*. 64(4): 487-535.

Heckman, James, and Edward Vytlacil. 2001. "Policy-Relevant Treatment Effects." *American Economic Review*. 91(2): 107-111.

Hotz, V. Joseph, Guido Imbens and Jacob Klerman. 2000. "The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program." NBER Working Paper No. 8007.

Hui, Shek-Wai and Jeffrey Smith. 2002a. "The Determinants of Participation in Adult Education and Training in Canada." Unpublished manuscript, University of Maryland.

Hui, Shek-Wai and Jeffrey Smith. 2002b. "Issues in the Design of the Adult Education and Training Survey." Unpublished manuscript, University of Maryland.

Imbens, Guido, and Joshua Angrist. 1994. "Identificaiton and Estimation of Local Average Treatment Effects." *Econometrica*. 62(4): 467-476.

Lechner, Michael, and Ruth Miquel. 2002. "Identification of the Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions." Discussion Paper, Department of Economics, Universität St. Gallen.

Lee, Lung-Fei. 1983. "Generalized Econometric Models with Selectivity." *Econometrica* 51(2): 507-512.

Lillard, Lee A and Tan, Hong W. (1992): "Private Sector Training: Who Gets It and What Are Its Effects? " in Ronald Ehrenberg (ed.), Research in Labor Economics, 13, J.A.I. Press, 1-62.

Michalopoulos, Charles, David Card, Lisa Gennetian, Kristen Harknett and Philip Robins. 2000. *The Self-Sufficiency Project at 36 Months: Effects of a Financial Work Incentive on Employment and Income*. Ottawa: Social Research and Demonstration Corporation.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press.

Park, Norm, Bob Power, Craig Riddell and Ging Wong. 1996. “An Assessment of the Impact of Government-Sponsored Training.” *Canadian Journal of Economics*. 29(Special Issue, Part 1). S93-S98.

Puhani, Patrick. 2000. “The Heckman Correction for Sample Selection and Its Critique: A Short Survey.” *Journal of Economic Surveys*. 14: 53-68.

Riddell, Craig. 1991. “Evaluation of Manpower and Training Programs: The North American Experience.” In *The Evaluation of Manpower, Training and Social Programs: The State of a Complex Art*.” Paris, OECD. 43-72.

Rosenbaum, Paul and Rubin, Donald. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika*. 70(1): 41-55.

Smith, Jeffrey. 2000. “A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies.” *Swiss Journal of Economics and Statistics*. 136(3):1-22.

Smith, Jeffrey, and Arthur Sweetman. 2001. “Improving the Evaluation of Employment and Training Programs in Canada.” Unpublished manuscript, University of Maryland.

Smith, Jeffrey, and Petra Todd. 2002. “Does Matching Overcome LaLonde’s Critique of Non-experimental Estimators.” *Journal of Econometrics*. Forthcoming.

Trott, Charles and John Baj. 1993. “An Analysis of Repeating in JTPA in Illinois.” Report Prepared for the Illinois Department of Commerce and Community Affairs by Northern Illinois University.

Waite, Linda, and Maggie Gallagher. 2001. The Case for Marriage: *Why Married People are Happier, Healthier, and Better Off Financially*. New York: Broadway Books.

Warburton, William and Warburton, Rebecca. 2002. “Measuring the Performance of Government Training Programs.” C.D. Howe Institute Commentary No. 165.

Figure 1A: Distribution of Propensity of Any Training (Men)

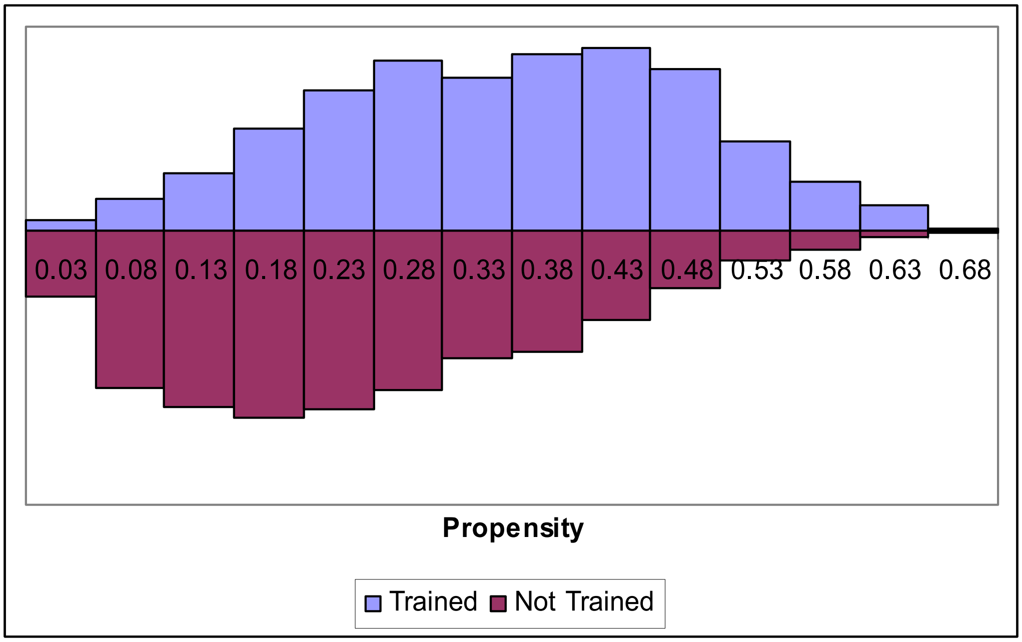

Figure 1B: Distribution of Propensity of Any Training (Women)

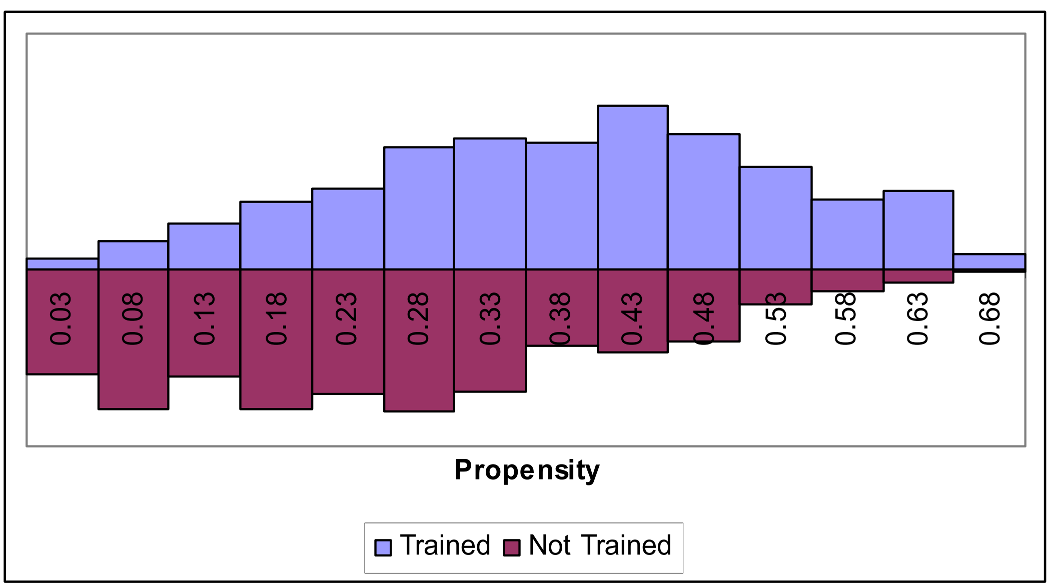

Figure 1C: Distribution of Propensity of Employer Financed Training (Men)

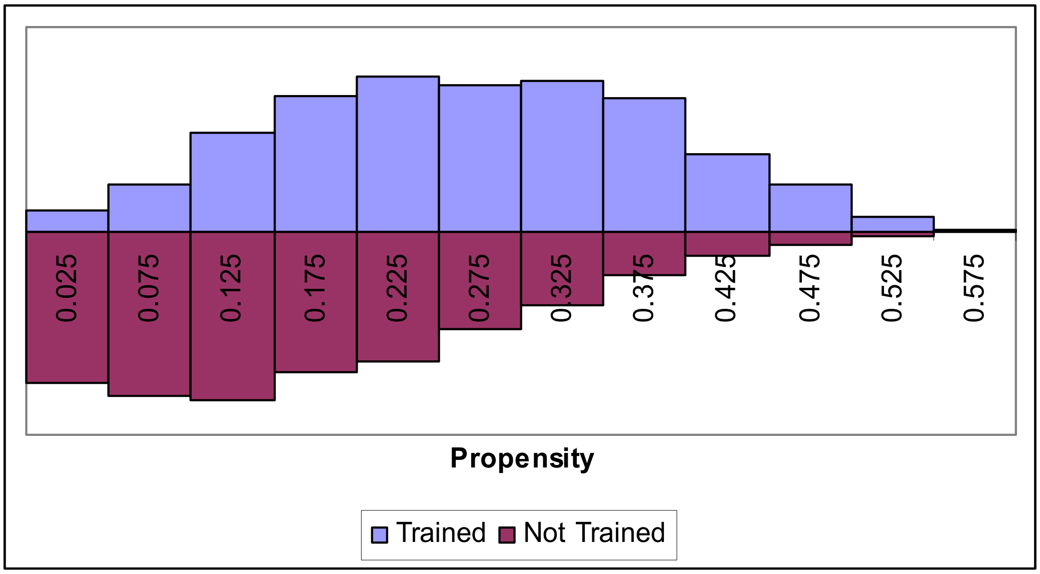

Figure 1D: Distribution of Propensity of Employer Financed Training (Women)

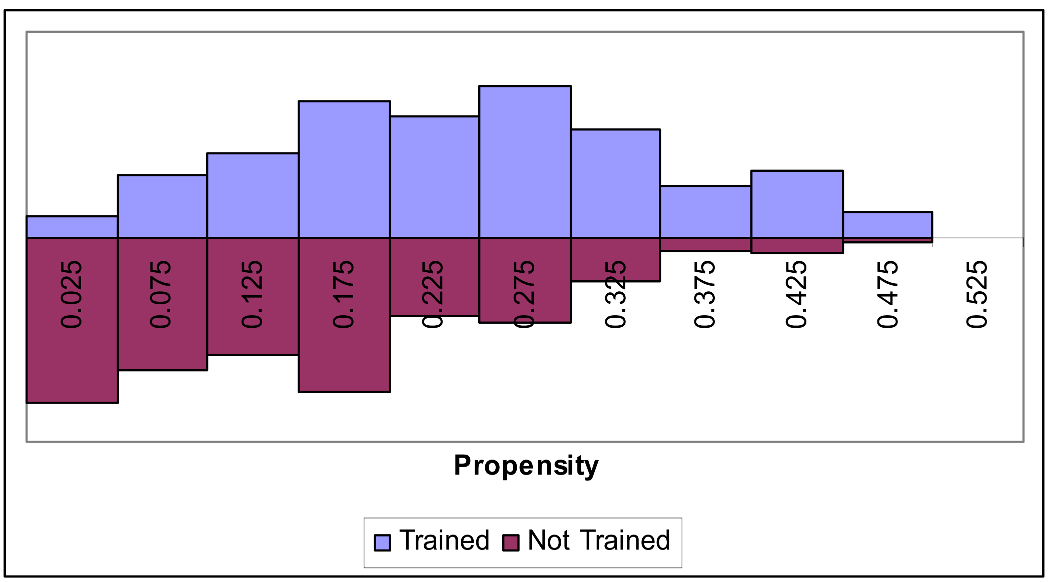

Figure 1E: Distribution of Propensity of Government Financed Training (Men)

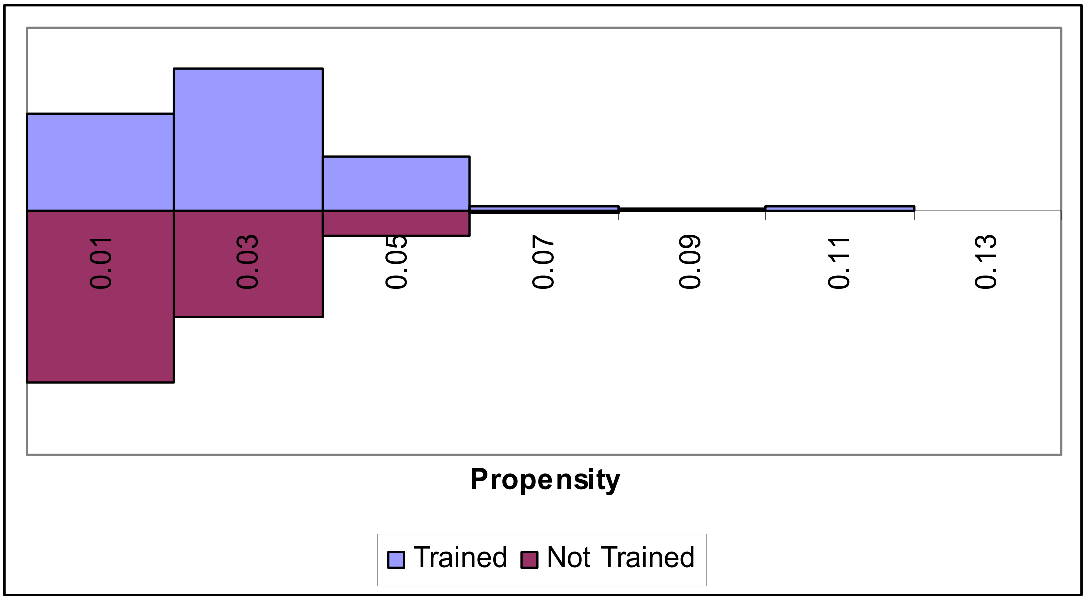

Figure 1F: Distribution of Propensity of Government Financed Training (Women)

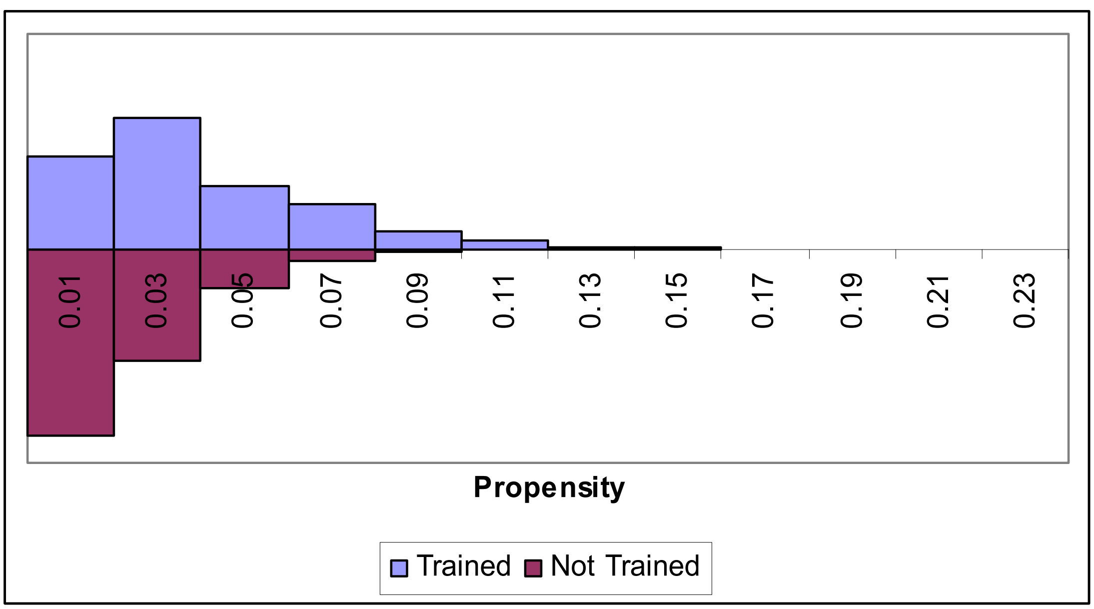

Figure 1G: Distribution of Propensity of Self Financed Training (Men)

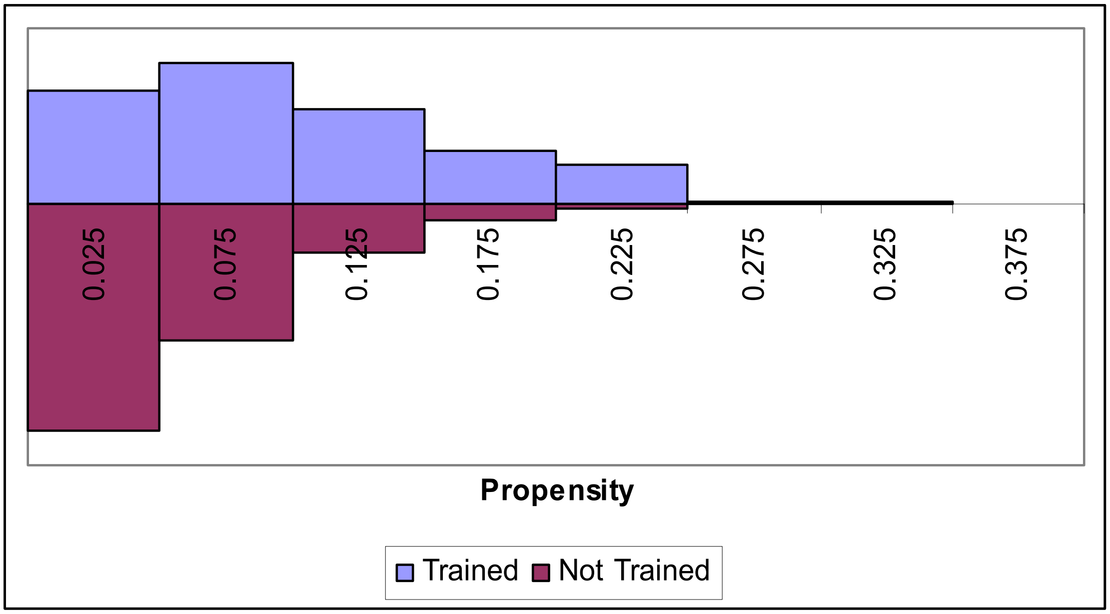

Figure 1H: Distribution of Propensity of Self Financed Training (Women)

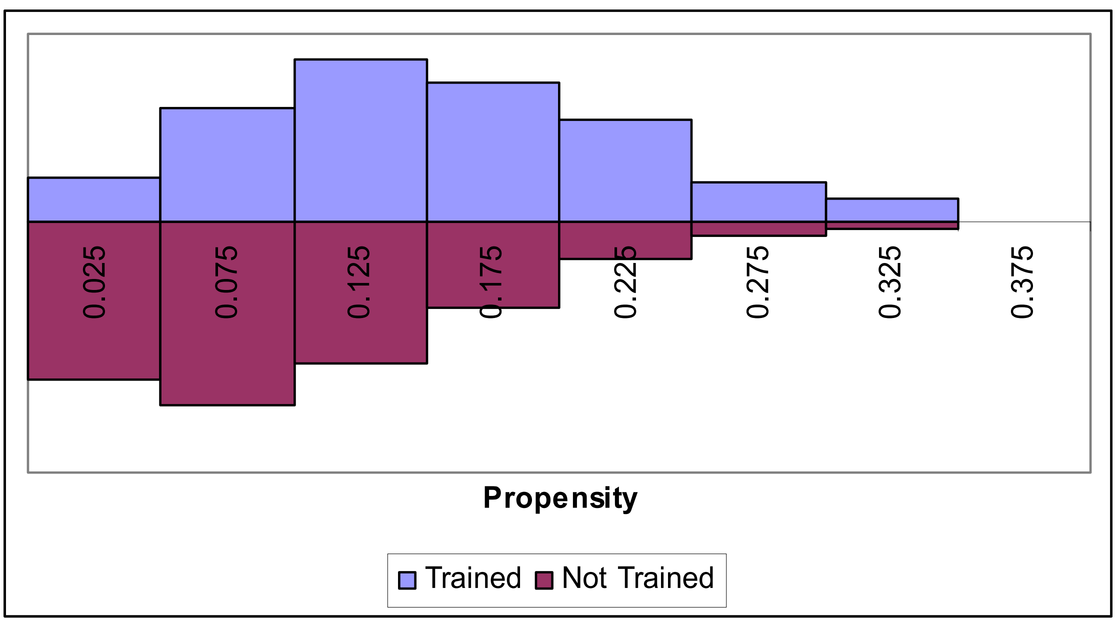

Table 1
AETS 1998 Total Observations and Sample Restrictions

	Male	Female
Total number of observations	14875	18535
Full Time Students In 1998	855	1107
Not Full Time Students:		
Age: 17-24	856	1052
Age: 65 and over	2416	3958
Sample used	10748	12418

Table 2
Summary Statistics on Training Incidence (Percentages)

	Men	Women
Program Participation		
Employer-Financed Programs	3.08	2.07
	(0.23)	(0.17)
Government-Financed Programs	0.84	1.21
	(0.14)	(0.15)
Self-Financed Programs	2.87	4.61
	(0.25)	(0.29)
Any Training Programs	6.83	7.91
	(0.36)	(0.36)
Course Participation		
Employer-Financed Courses	17.07	15.52
	(0.51)	(0.47)
Government-Financed Courses	1.31	1.28
	(0.19)	(0.14)
Self-Financed Courses	3.79	6.78
	(0.26)	(0.35)
Any Training Courses	21.90	22.93
	(0.57)	(0.55)
Any Training Participation		
Employer-Financed Training	19.39	17.00
	(0.53)	(0.48)
Government-Financed Training	2.11	2.40
	(0.23)	(0.20)
Self-Financed Training	6.38	10.81
	(0.35)	(0.43)
Any Training	27.07	28.76
	(0.62)	(0.60)

Notes:

Estimated standard errors in parentheses. The sample sizes are 10,748 and 12,418 for men and women, respectively.

Table 3
Summary Statistics on Total Training Time in Hours

Any Training		QUARTILES		
MEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	69.1	12	24	48
Government-Financed Training	591.5	90	425	1026
Self-Financed Training	526.2	32	304	960
Any Training	320.7	16	48	425
WOMEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	57.7	8	18	42
Government-Financed Training	605.3	99	480	1000
Self-Financed Training	443.4	24	180	745
Any Training	314.7	12	50	450

Notes:
Statistics use lower bounds for spells in progress at the time of the AETS interview and for top coded spells. See the discussion in the text. The sample includes respondents in the sample defined in Table 1 with positive reported training time.

Table 3 (continued)

Summary Statistics on Total Training Time in Hours

Training Programs			QUARTILES	
MEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	252.5	40	90	240
Government-Financed Training	770.4	320	736	1170
Self-Financed Training	730.6	225	600	1120
Any Training	676.1	156	480	1080
WOMEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	219.2	39	96	280
Government-Financed Training	783.5	352	720	1200
Self-Financed Training	646.4	180	496	990
Any Training	624.6	150	480	960

Training Courses			QUARTILES	
MEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	34.9	10	18	38
Government-Financed Training	75.3	6	19	50
Self-Financed Training	66.9	6	18	40
Any Training	45.6	7	18	42
WOMEN	MEAN	1st-quartile	Median	3rd-quartile
Employer-Financed Training	34.7	6	18	30
Government-Financed Training	117.8	10	30	112
Self-Financed Training	62.7	6	16	35
Any Training	50.6	6	18	36

Notes:

Statistics use lower bounds for spells in progress at the time of the AETS interview and for top coded spells. See the discussion in the text. The sample includes respondents in the sample defined in Table 1 with positive reported training time.

Table 4
Summary Statistics for Outcome Variables Used in the Analysis

	Weekly Earnings ($)		Employment (%)	
	Men	Women	Men	Women
Overall	477.54 (6.63) *464.36*	302.84 (4.73) *338.76*	80.01 (0.55)	67.32 (0.60)
Not in any training	413.91 (7.38) *443.17*	236.25 (5.16) *300.74*	76.07 (0.68)	60.00 (0.74)
In Any Training	648.96 (13.28) *476.68*	467.83 (9.35) *369.90*	90.62 (0.82)	85.46 (0.84)
Employer-Financed Training	748.02 (14.68) *462.30*	582.70 (10.35) *334.91*	96.31 (0.49)	96.58 (0.52)
Government-Financed Training	331.63 (51.10) *423.39*	186.13 (23.74) *284.93*	61.26 (5.52)	48.84 (4.23)
Self-Financed Training	448.53 (23.39) *417.96*	367.75 (16.81) *377.87*	83.44 (2.10)	77.36 (1.62)

Notes:
Estimated standard errors appear in parentheses, and standard deviations appear in italics. Outcomes are measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. Training refers to training in calendar year 1997.

Table 5
Summary Statistics on Charateristics of Individuals and their Jobs

	Men	Women
Province		
Newfoundland	1.91 (0.10)	1.89 (0.09)
PEI	0.44 (0.03)	0.44 (0.02)
Nova Scotia	2.99 (0.13)	3.13 (0.13)
New Brunswick	2.51 (0.11)	2.53 (0.10)
Quebec	24.91 (0.62)	24.92 (0.58)
Ontario	38.12 (0.72)	38.29 (0.68)
Manitoba	3.51 (0.15)	3.47 (0.14)
Saskatchewan	2.96 (0.13)	2.97 (0.12)
Alberta	9.49 (0.36)	9.20 (0.33)
B.C.	13.16 (0.46)	13.16 (0.43)
Regions		
Census Metro Area	66.01 (0.58)	64.96 (0.56)
Toronto/Vancouver/Montreal	35.05 (0.77)	35.54 (0.72)
Urban Centre	7.25 (0.27)	7.98 (0.29)
Rural Area	16.07 (0.40)	15.96 (0.38)

Age Group	Men	Women
25-34	27.55 (0.64)	27.22 (0.59)
35-44	31.52 (0.63)	31.50 (0.60)
45-54	24.92 (0.60)	24.87 (0.58)
55-64	16.01 (0.50)	16.41 (0.49)
Age (Years)	42.40 (0.15)	42.50 (0.14)
Age Squared	1907.22 (12.82)	1917.11 (12.23)
Level of Education		
Some Secondary	13.82 (0.48)	13.00 (0.40)
Grade 11-13 graduate	18.79 (0.56)	21.75 (0.56)
Some Post-Secondary	7.38 (0.36)	8.19 (0.35)
Post-secondary Certificate or Diplomas	32.49 (0.63)	32.75 (0.61)
Bachelor, Master or PhD	20.11 (0.58)	17.10 (0.51)
Presence of Spouse	72.09 (0.64)	71.68 (0.57)
Spouse's Level of Education		
Some Secondary	9.25 (0.39)	9.70 (0.36)
Grade 11-13 graduate	16.73 (0.52)	13.45 (0.46)
Some Post-Secondary	6.05 (0.33)	4.76 (0.27)
Post-secondary Certificate or Diplomas	23.93 (0.58)	24.17 (0.56)
Bachelor, Master or PhD	39.84 (0.69)	42.18 (0.65)

	Men	Women
Number of Children (age below 18)		
1 Child	17.83 (0.53)	19.37 (0.53)
2 or More Children	31.87 (0.66)	34.95 (0.63)
Number of Preschool Children		
1 Pre-school Child	11.86 (0.43)	13.24 (0.43)
2 or More Pre-school Children	6.00 (0.32)	6.22 (0.31)
Spouse & Children interactions		
Presence of Spouse and Children	47.59 (0.70)	45.41 (0.66)
Presence of Spouse and 2 or More Children	31.10 (0.65)	30.56 (0.62)
Presence of Spouse and Preschool Children	17.64 (0.52)	17.16 (0.49)
Other Personal Characteristics		
Born Foreign	19.68 (0.65)	20.55 (0.64)
Job Characteristics		
Employed in 97	87.11 (0.47)	73.50 (0.57)
Full Time Worker in 97	72.85 (0.61)	45.66 (0.66)
Professional, Managerial or Administrative Occupation	28.30 (0.63)	30.41 (0.61)
Blue Collar Occupation	38.43 (0.67)	8.35 (0.42)
Union Member	26.35 (0.60)	21.63 (0.54)

Job Characteristics	Men	Women
Firm Size: Less than 20	29.43 (0.62)	23.59 (0.54)
20 - 99	13.00 (0.49)	9.93 (0.39)
100 - 199	5.59 (0.32)	4.41 (0.28)
200 - 499	6.21 (0.34)	5.91 (0.32)
500 or over	31.18 (0.65)	27.71 (0.60)
Missing Indicators		
Missing Place of Birth	0.52 (0.12)	0.37 (0.09)
Missing Union Status	0.11 (0.03)	0.07 (0.02)
Missing Firm Size	1.58 (0.18)	1.81 (0.19)
Missing Occupation	0.46 (0.13)	0.08 (0.03)

Notes:
Estimated standard errors appear in parentheses. Units are percentages for binary variables and means for continuous variables

Table 6
Definitions of Instruments and Conditioning Sets

<u>List of Set A Variables</u>

Dummies for Provinces and Regions of Residence

Newfoundland	PEI
Nova Scotia	New Brunswick
Quebec	Manitoba
Saskatchewan	Alberta
B.C.	
Census Metro Area	Toronto/Montreal/Vancouver
Urban Centres	Rural Area

Personal Characteristics

Age	Age Squared
Some Secondary	Grade 11-13
Some Post-secondary	Post-secondary Certificates of Diplomas
Bachelor/Master/PhD	
Foreign Born	Missing Place of Birth

<u>List of Set B Variables</u>

Characteristics of the job in 1997

Employment Status in 1997	Full-time Working in 97
Union Member	Blue Collar Occupation
Firm Size: 20-99	Professional, Managerial or Adminstrative Occupation
Firm Size: 200-499	
Firm Size: 100-199	
Firm Size: 500 or over	Missing Union Information
Missing Occupation Information	Missing Firm Size Information

<u>List of Set C Variables</u>

Family Characteristics

Presence of Spouse	
Spouse: Some Secondary Education	Spouse: Grade 11-13
Spouse: Some Post-secondary	Spouse: Post-Secondary Certificate or Diplomas
Spouse: Bachelor/Master/PhD	
1 Child	2 or More Children
1 Child and Spouse Present	2 or More Children and Spouse Present
1 Preschool Child	2 or More Preschool Children
Preschool Children and Spouse Present	

Table 7
Estimated Impacts of Any Training on Employment

Impacts on Employment (%)	Men Any Training	Women Any Training
Linear Probability Model	7.82 (1.04)***	15.91 (1.23)***
Probit Marginal Effect	9.15 (1.18)***	18.60 (1.38)***
Nearest Neighbour Matching (Set A)		
1 to 1	7.89 (0.97)***	17.06 (1.08)***
1 to 2	7.72 (0.81)***	16.77 (0.91)***
1 to 5	9.01 (0.71)***	16.63 (0.78)***
Nearest Neighbour Matching (Sets A and B)		
1 to 1	2.66 (0.89)***	2.59 (0.96)***
1 to 2	2.99 (0.78)***	2.78 (0.87)***
1 to 5	2.49 (0.70)***	2.87 (0.80)***
Instrumental Variable	65.56 (12.74)***	163.06 (34.59)***
Bivariate Probit Marginal Effect	66.53 (0.29)***	69.04 (0.23)

Notes:

Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes are 10,469 for men and 11,938 for women. Employment is measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable and bivariate probit models condition on covariate set A and use covariate set C as instruments.

Table 8A
Estimated Impacts of Any Training on Employment by Type - Men

Impacts on Employment (%)	Employer-Financed	Government-Financed	Self-Financed
Linear Probability Model	12.49	-16.14	-0.35
	(0.81)***	(5.12)***	(2.48)
Probit Marginal Effect	31.28	-19.30	4.30
	(1.07)***	(4.60)***	(2.24)*
Nearest Neighbour Matching (Set A)			
1 to 1	11.49	-18.00	-3.61
	(0.98)***	(4.38)***	(2.30)
1 to 2	12.15	-17.75	-3.61
	(0.80)***	(3.98)***	(2.04)*
1 to 5	12.79	-14.3	-1.45
	(0.67)***	(3.70)***	(1.91)
Nearest Neighbour Matching (Sets A and B)			
1 to 1	0.94	-5.50	2.61
	(0.73)	(4.73)	(2.48)
1 to 2	0.52	-3.75	1.81
	(0.62)	(4.35)	(2.11)
1 to 5	1.00	-1.4	3.17
	(0.57)*	(3.95)	(1.89)*
Instrumental Variable	70.61	-40.98	-132.70
	(16.65)***	(141.07)	(54.02)**

Notes:
Estimated (robust) standard errors appear in parentheses. The “*” denotes statistical significance at the ten percent level, “**” at the five percent level, and “***” at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes for the linear probability model and the IV models are 10,469. The sample sizes for the nearest neighbor matching are 10,611, 10,720 and 10,631 for employer-financed, government-financed and self-financed training, respectively. Employment is measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable and bivariate probit models condition on covariate set A and use covariate set C as instruments.

Table 8B
Estimated Impacts of Any Training on Employment by Type - Women

Impacts on Employment (%)	Employer-Financed	Government-Financed	Self-Financed
Linear Probability Model	25.24	-17.91	4.09
	(0.98)***	(4.33)***	(1.96)**
Probit Marginal Effect	30.76	-20.32	3.57
	(1.17)***	(5.16)***	(2.24)
Nearest Neighbour Matching (Set A)			
1 to 1	26.29	-24.33	1.13
	(1.15)***	(4.10)***	(2.00)
1 to 2	25.92	-23.95	1.65
	(0.89)***	(3.71)***	(1.77)
1 to 5	27.29	-22.74	1.79
	(0.67)***	(3.36)***	(1.59)
Nearest Neighbour Matching (Sets A and B)			
1 to 1	2.77	-6.84	3.40
	(0.72)***	(4.35)	(2.02)*
1 to 2	2.90	-4.75	3.76
	(0.61)***	(3.93)	(1.81)**
1 to 5	3.20	-6.62	3.21
	(0.53)***	(3.60)*	(1.61)**
Instrumental Variable	188.60	-72.77	156.43
	(64.36)***	(110.33)	(106.65)

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes for the linear probability and IV models are 11,938. The sample sizes for the nearest neighbor matching are 12,255, 12,370 and 12,145 for employer-financed, government-financed and self-financed training, respectively. Employment is measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable and bivariate probit models condition on covariate set A and use covariate set C as instruments.

Table 9
Estimated Impacts of Any Training on Usual Weekly Earnings

Impacts on Weekly Earnings ($)	Men Any Training	Women Any Training
Linear Model	172.8 (15.8)***	158.9 (11.4)***
Nearest Neighbour Matching (Set A)		
1 to 1	172.1 (12.5)***	154.8 (8.84)***
1 to 2	161.6 (11.1)***	153.2 (7.9)***
1 to 5	162.6 (10.1)***	151.6 (7.2)***
Nearest Neighbour Matching (Sets A and B)		
1 to 1	33.4 (12.9)***	28.1 (9.1)***
1 to 2	45.2 (11.5)***	20.4 (8.5)***
1 to 5	44.5 (10.6)***	22.0 (7.9)***
Instrumental Variable	1,086.5 (188.4)***	909.5 (205.3)***
Bivariate Normal Model	407.2 (90.4)***	204.4 (18.5)***

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes are 10,469 for men and 11,938 for women. Usual weekly earnings are measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable and bivariate normal models condition on covariate set A and use covariate set C as instruments.

Table 10A
Estimated Impacts of Any Training on Usual Weekly Earnings by Type - Men

Impacts on Weekly Earnings ($)	Employer-Financed	Government-Financed	Self-Financed
Linear Model	256.6	-122.4	-60.5
	(17.1)***	(50.5)**	(28.4)**
Nearest Neighbour Matching (Set A)			
1 to 1	251.5	-106.7	-94.9
	(14.4)***	(41.2)***	(27.5)***
1 to 2	248.3	-113.3	-99.5
	(12.4)***	(34.9)**	(24.2)***
1 to 5	243.4	-104.7	-88.0
	(11.2)***	(30.9)***	(21.6)***
Nearest Neighbour Matching (Sets A and B)			
1 to 1	59.4	-6.3	-50.1
	(14.2)***	(39.8)	(27.3)*
1 to 2	50.2	-20.3	-82.8
	(12.9)***	(35.4)	(23.8)***
1 to 5	60.6	16.2	-74.3
	(11.8)***	(31.7)	(21.1)***
Instrumental Variable	1,162.5	720.1	-969.6
	(203.6)***	(1,736.9)	(646.0)

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes for the linear probability model and the IV models are 10,469. The sample sizes for the nearest neighbor matching are 10,611, 10,720 and 10,631 for employer-financed, government-financed and self-financed training, respectively. Usual weekly earnings are measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable model conditions on covariate set A and uses covariate set C as instruments.

Table 10B
Estimated Impacts of Any Training on Usual Weekly Earnings by Type - Women

Impacts on Weekly Earnings ($)	Employer-Financed	Government-Financed	Self-Financed
Linear Model	256.2	-93.0	15.1
	(11.9)***	(23.3)***	(19.2)
Nearest Neighbour Matching (Set A)			
1 to 1	261.9	-139.4	-33.8
	(10.5)***	(25.4)***	(15.8)**
1 to 2	248.7	-116.1	-17.8
	(9.4)***	(21.3)***	(13.8)
1 to 5	253.3	-116.4	-17.5
	(8.4)***	(19.3)***	(12.5)
Nearest Neighbour Matching (Sets A and B)			
1 to 1	40.6	-19.5	3.0
	(10.8)***	(23.9)	(15.3)
1 to 2	47.2	-31.4	-4.9
	(10.0)***	(22.1)	(13.8)
1 to 5	48.4	-37.3	-18.7
	(9.3)***	(20.0)*	(12.7)
Instrumental Variable	1,112.1	-666.9	1,035.2
	(391.0)***	(681.5)	(679.1)

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The impact estimates are for training received in calendar year 1997. Respondents whose training was in progress at the time of the AETS interview are not included in the estimation. The sample sizes for the linear probability model and the IV models are 11,938. The sample sizes for the nearest neighbor matching are 12,255, 12,370 and 12,145 for employer-financed, government-financed and self-financed training, respectively. Usual weekly earnings are measured in March 1998 for respondents in Quebec and in January 1998 for all other respondents. The linear probability model includes covariate sets A and C. Propensity scores for the nearest neighbor matching estimator include either covariate set A or covariate sets A and B. The instrumental variable model conditions on covariate set A and uses covariate set C as instruments.

Table A1
Complete Estimates for Linear Probability, Probit and IV Models in Table 7
Estimated Coefficients (x100) and Probit Marginal Effects (x100)

	LPM		Probit		IV	
	Men	Women	Men	Women	Men	Women
Newfoundland	-23.52	-15.80	-26.12	-17.45	-17.29	-1.17
	(2.59)***	(2.52)***	(3.15)***	(2.95)***	(3.28)***	(5.15)
PEI	-12.55	-5.07	-13.61	-5.28	-7.88	5.67
	(3.06)***	(2.82)*	(3.44)***	(3.20)*	(3.60)**	(5.02)
Nova Scotia	-7.42	-6.05	-8.42	-6.92	-7.52	-0.91
	(1.99)***	(2.05)***	(2.35)***	(2.41)***	(2.44)***	(3.70)
New Brunswick	-11.35	-10.46	-12.46	-11.95	-6.85	-0.71
	(2.11)***	(2.10)***	(2.49)***	(2.46)***	(2.60)***	(4.11)
Quebec	-5.27	-2.89	-6.01	-3.36	1.39	17.38
	(1.50)***	(1.63)*	(1.70)***	(1.87)*	(2.32)	(5.43)***
Manitoba	7.03	2.79	7.19	2.95	8.28	3.51
	(1.59)***	(1.99)	(1.47)***	(2.20)	(2.04)***	(3.59)
Saskatchewan	4.81	4.60	4.78	5.18	5.60	4.61
	(1.74)***	(1.91)**	(1.72)***	(2.14)**	(2.27)***	(3.54)
Alberta	5.33	4.54	6.06	5.11	7.36	5.17
	(1.49)***	(1.86)**	(1.57)***	(2.08)**	(2.09)***	(3.34)
B.C.	-6.19	-4.62	-7.01	-5.48	-5.97	-5.88
	(1.77)***	(1.95)**	(2.14)***	(2.31)**	(2.13)***	(3.56)*
Census Metro Area	-1.24	0.91	-1.45	0.94	-2.03	1.22
	(1.39)	(1.58)	(1.42)	(1.75)	(1.78)	(2.64)
Toronto/Montreal/Vancouver	4.01	4.71	4.35	5.51	7.18	2.85
	(1.44)***	(1.60)***	(1.51)***	(1.79)***	(1.94)***	(2.94)
Urban Centres	-1.56	1.91	-1.62	2.09	-2.80	-1.22
	(1.98)	(1.99)	(2.05)	(2.16)	(2.31)	(3.51)
Rural Area	-2.67	-1.56	-2.94	-1.80	-3.03	-0.18
	(1.58)*	(1.70)	(1.63)*	(1.88)	(1.83)*	(2.83)
Age	5.60	5.00	4.77	5.20	4.94	2.84
	(0.46)***	(0.44)***	(0.46)***	(0.51)***	(0.66)***	(1.08)***
Age Squared	-0.07	-0.06	-0.06	-0.07	-0.06	-0.03
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***
Some Secondary	13.08	3.75	8.76	3.15	14.12	6.43
	(2.93)***	(2.94)	(2.66)***	(3.04)	(3.13)***	(3.75)*
Grade 11-13	8.71	13.03	8.38	12.66	4.96	-0.26
	(2.16)***	(2.09)***	(2.32)***	(2.19)***	(2.61)*	(4.34)
Some Post-Secondary	-1.82	3.44	-1.79	3.15	-7.29	-5.74
	(2.28)	(2.53)	(2.36)	(2.70)	(3.19)**	(4.99)
Certificate or Diploma	2.80	5.20	3.04	6.17	2.27	-4.84
	(2.07)	(2.33)**	(2.18)	(2.55)**	(2.71)	(4.90)
Bachelor/Master/PhD	6.16	3.91	7.40	5.76	-0.85	-15.51
	(1.22)***	(1.55)**	(1.39)***	(2.03)***	(2.39)	(5.64)***
Foreign Born	-2.22	-7.06	-3.25	-8.54	-0.15	6.31
	(1.61)	(1.83)***	(1.86)*	(2.12)***	(2.07)	(4.78)
Missing Place of Birth	-27.03	4.70	-33.35	7.51	-31.12	15.22
	(10.40)***	(5.86)	(11.08)***	(7.53)	(11.65)***	(18.12)
Any Training	7.82	15.91	9.15	18.60	65.56	163.06
	(1.04)***	(1.23)***	(1.18)***	(1.38)***	(12.74)***	(34.59)***
Observations	10,469	11,938	10,469	11,938	10,469	11,938
Log-Likelihood			-4414.395	-6541.038		
R-squared	0.17	0.17				
Observed Probability (at mean)			0.80	0.67		
Predicted Probability (at mean)			0.83	0.69		

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The estimated constant term is omitted.

Table A2
Complete Estimates for Bivariate Probit Model in Table 7
Estimated Probit Marginal Effects (x100)

	Bi-Probit			
	Employment		Training Incidence	
	Men	Women	Men	Women
Newfoundland	-0.46 (0.09)***	-0.29 (0.07)***	-0.37 (0.09)***	-0.35 (0.08)***
PEI	-0.23 (0.10)**	-0.04 (0.08)	-0.25 (0.10)**	-0.23 (0.09)***
Nova Scotia	-0.23 (0.07)***	-0.12 (0.06)**	0.02 (0.07)	-0.10 (0.07)
New Brunswick	-0.21 (0.07)***	-0.20 (0.06)***	-0.22 (0.07)***	-0.20 (0.07)***
Quebec	0.01 (0.06)	0.08 (0.06)	-0.38 (0.06)***	-0.45 (0.06)***
Manitoba	0.30 (0.07)***	0.08 (0.06)	-0.06 (0.07)	-0.01 (0.07)
Saskatchewan	0.18 (0.07)**	0.13 (0.06)**	-0.06 (0.07)	0.02 (0.06)
Alberta	0.26 (0.07)***	0.13 (0.06)**	-0.11 (0.06)*	-0.02 (0.06)
B.C.	-0.20 (0.07)***	-0.14 (0.06)**	0.01 (0.06)	0.01 (0.06)
Census Metro Area	-0.07 (0.05)	0.03 (0.05)	0.05 (0.05)	-0.01 (0.05)
Toronto/Montreal/Vancouver	0.23 (0.06)***	0.12 (0.05)**	-0.15 (0.06)**	0.04 (0.06)
Urban Centres	-0.08 (0.07)	0.03 (0.06)	0.05 (0.07)	0.08 (0.07)
Rural Area	-0.10 (0.05)*	-0.03 (0.05)	0.01 (0.06)	-0.01 (0.06)
Age	0.14 (0.02)***	0.11 (0.02)***	0.00 (0.02)	0.05 (0.02)***
Age Squared	0.00 (0.00)***	0.00 (0.00)***	0.00 (0.00)	0.00 (0.00)***
Some Secondary	0.31 (0.08)***	0.12 (0.08)	-0.05 (0.11)	0.12 (0.15)
Grade 11-13	0.13 (0.08)*	0.20 (0.06)***	0.26 (0.07)***	0.45 (0.07)***
Some Post-Secondary	-0.21 (0.09)**	0.00 (0.08)	0.28 (0.08)***	0.17 (0.08)**
Certificate or Diploma	0.07 (0.08)	0.07 (0.07)	0.03 (0.08)	0.22 (0.07)***
Bachelor/Master/PhD	0.07 (0.07)	-0.01 (0.06)	0.30 (0.06)***	0.33 (0.06)***

Table A2 (continued)
Complete Estimates for Bivariate Probit Model in Table 7
Estimated Probit Marginal Effects (x100)

	Bi-Probit			
	Employment		Training Incidence	
	Men	Women	Men	Women
Foreign Born	-0.04	-0.09	-0.16	-0.30
	(0.06)	(0.06)	(0.06)**	(0.06)***
Missing Place of Birth	-0.85	0.27	0.21	-0.18
	(0.27)***	(0.27)	(0.30)	(0.33)
Spouse Present			0.44	0.07
			(0.07)***	(0.07)
Spouse - Some Secondary			0.08	0.19
			(0.16)	(0.11)*
Spouse - 11-13			0.13	0.02
			(0.08)*	(0.08)
Spouse - Some Post-Secondary			0.03	-0.01
			(0.08)	(0.10)
Spouse - Certificate or Diploma			-0.01	0.06
			(0.08)	(0.09)
Spouse - Bachelors or More			0.15	0.02
			(0.06)**	(0.06)
1 Child			0.13	-0.04
			(0.11)	(0.09)
2 or More Children			0.20	-0.08
			(0.19)	(0.10)
1 Child with Spouse			-0.04	-0.03
			(0.13)	(0.11)
2 or More Children with Spouse			-0.14	0.05
			(0.20)	(0.12)
1 Preschool Child			-0.22	-0.32
			(0.24)	(0.11)***
2 or More Preschool Children			-0.16	-0.27
			(0.08)*	(0.08)***
Preschool Children with Spouse			0.09	0.13
			(0.24)	(0.12)
Any Training	1.67	1.64		
	(0.09)***	(0.12)***		
Observations	11,469	11,938		
Log-Likelihood	-7207403	-8007788		
Rho	-0.85	-0.71		
Pr(Rho=0)	0.00	0.00		

Notes:
Estimated (robust) standard errors appear in parentheses. The “*” denotes statistical significance at the ten percent level, “**” at the five percent level, and “***” at the one percent level. The estimated constant term is omitted.

Table A3

Complete Estimates for Linear Probability, Probit and IV Models in Tables 8A and 8B

Estimated Coefficients (x100) and Probit Marginal Effects (x100)

	LPM		Probit		IV	
	Men	Women	Men	Women	Men	Women
Newfoundland	-23.00	-15.25	-25.12	-16.66	-18.71	2.02
	(2.58)***	(2.53)***	(3.17)***	(3.02)***	(3.82)***	(6.44)
PEI	-11.75	-4.99	-11.98	-5.14	-5.33	6.03
	(3.05)***	(2.81)*	(3.35)***	(3.23)	(4.37)	(5.53)
Nova Scotia	-7.67	-5.83	-8.85	-6.77	-6.75	2.58
	(1.99)***	(2.05)***	(2.41)***	(2.46)***	(3.56)*	(4.84)
New Brunswick	-10.89	-10.08	-11.41	-11.44	-7.30	2.61
	(2.10)***	(2.07)***	(2.45)***	(2.50)***	(3.25)**	(4.75)
Quebec	-5.14	-2.47	-5.68	-2.94	-0.42	21.23
	(1.48)***	(1.62)	(1.64)***	(1.85)	(3.54)	(5.92)***
Manitoba	6.95	3.20	6.85	3.64	9.92	5.18
	(1.59)***	(1.98)	(1.40)***	(2.14)*	(3.51)***	(4.65)
Saskatchewan	4.47	4.70	4.39	5.54	2.81	7.04
	(1.73)***	(1.88)**	(1.68)***	(2.07)***	(3.12)	(4.26)*
Alberta	5.20	4.59	5.68	5.09	6.48	6.73
	(1.49)***	(1.84)**	(1.52)***	(2.05)**	(2.72)**	(3.74)*
B.C.	-5.96	-4.14	-6.56	-4.90	-1.45	-3.47
	(1.75)***	(1.93)**	(2.07)***	(2.29)**	(3.08)	(4.02)
Census Metro Area	-1.05	0.92	-1.10	1.03	-0.10	1.98
	(1.38)	(1.57)	(1.38)	(1.76)	(2.18)	(3.12)
Toronto/Montreal/Vancouver	4.20	4.59	4.50	5.56	6.14	2.46
	(1.42)***	(1.59)***	(1.42)***	(1.77)***	(2.41)**	(3.30)
Urban Centres	-1.33	2.41	-1.42	2.62	0.89	2.25
	(1.97)	(2.00)	(2.00)	(2.15)	(2.95)	(4.31)
Rural Area	-2.57	-1.63	-2.80	-1.90	-2.57	0.73
	(1.58)	(1.69)	(1.60)*	(1.90)	(2.28)	(3.61)
Age	5.35	4.69	4.28	4.71	2.25	1.66
	(0.45)***	(0.44)***	(0.43)***	(0.49)***	(1.06)**	(1.85)
Age Squared	-0.07	-0.06	-0.06	-0.06	-0.04	-0.02
	(0.01)***	(0.01)***	(0.00)***	(0.01)***	(0.01)***	(0.02)
Some Secondary	13.19	4.16	8.53	3.49	13.17	8.62
	(2.92)***	(2.92)	(2.60)***	(2.97)	(3.70)***	(3.88)**
Grade 11-13	8.11	12.52	7.34	11.54	2.87	-2.70
	(2.11)***	(2.06)***	(2.16)***	(2.15)***	(2.98)	(4.80)
Some Post-Secondary	-1.75	3.71	-1.69	3.61	-0.57	-4.26
	(2.24)	(2.44)	(2.26)	(2.62)	(3.74)	(5.08)
Certificate or Diploma	2.72	4.97	2.94	5.75	0.33	-8.10
	(2.03)	(2.25)**	(2.09)	(2.48)**	(3.08)	(5.77)
Bachelor/Master/PhD	6.03	2.98	7.20	4.57	7.85	-22.07
	(1.21)***	(1.55)*	(1.34)***	(2.08)**	(3.90)**	(6.86)***
Foreign Born	-2.09	-5.90	-3.09	-7.00	1.37	12.05
	(1.59)	(1.79)***	(1.78)*	(2.06)***	(2.67)	(6.29)*
Missing Place of Birth	-25.29	4.12	-31.70	6.85	-32.17	12.08
	(10.66)**	(6.08)	(12.11)***	(7.58)	(14.33)**	(21.08)
Employer-Financed Training	12.49	25.24	15.31	31.28	70.61	188.60
	(0.81)***	(0.98)***	(0.89)***	(1.07)***	(16.65)***	(64.36)***
Government-Financed Training	-16.14	-17.91	-17.43	-19.30	-40.98	-72.77
	(5.12)***	(4.33)***	(5.75)***	(4.60)***	(141.07)	(110.33)
Self-Financed Training	-0.35	4.09	-1.44	4.30	-132.70	156.43
	(2.48)	(1.96)**	(2.71)	(2.24)*	(54.02)**	(106.65)
Observations	10,469	11,938	10,469	11,938	10,469	11,938
Log-Likelihood			-4317.15	-6314.39		
R-squared	0.18	0.19				
Observed Probability			0.80	0.67		
Predicted Probability (at mean)			0.84	0.71		

Notes:

Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The estimated constant term is omitted.

Table A4
Complete Estimates for Linear and IV Models in Table 9
Estimated Coefficients (x100)

	Linear Model		Instrumental Variables	
	Men	Women	Men	Women
Newfoundland	-194.2	-108.6	-95.6	-33.9
	(22.9)***	(12.3)***	(36.2)***	(27.8)
PEI	-185.1	-39.9	-111.2	14.9
	(22.7)***	(15.8)**	(38.0)***	(27.2)
Nova Scotia	-125.4	-70.5	-127.1	-44.3
	(19.5)***	(12.9)***	(29.8)***	(20.5)**
New Brunswick	-90.2	-53.4	-19.0	-3.7
	(19.7)***	(13.1)***	(30.2)	(23.4)
Quebec	-65.8	-52.8	39.5	50.6
	(18.2)***	(12.1)***	(30.8)	(32.0)
Manitoba	-45.9	-58.5	-26.2	-54.8
	(19.3)**	(12.2)***	(28.4)	(19.6)***
Saskatchewan	-97.1	-51.6	-84.5	-51.6
	(19.7)***	(12.3)***	(28.2)***	(20.0)***
Alberta	-31.6	-35.7	0.6	-32.5
	(20.7)	(12.9)***	(29.0)	(19.4)*
B.C.	-55.9	-40.5	-52.3	-47.0
	(21.2)***	(15.2)***	(28.0)*	(21.0)**
Census Metro Area	46.8	28.4	34.4	30.0
	(16.1)***	(10.6)***	(22.1)	(15.4)*
Toronto/Montreal/Vancouver	-1.4	60.8	48.8	51.3
	(19.1)	(12.8)***	(27.3)*	(17.6)***
Urban Centres	-4.8	12.0	-24.4	-4.0
	(21.1)	(13.5)	(29.1)	(20.6)
Rural Area	-91.9	-48.4	-97.5	-41.4
	(16.3)***	(10.5)***	(22.1)***	(16.0)***
Age	63.0	34.2	52.5	23.2
	(4.5)***	(3.2)***	(6.6)***	(5.8)***
Age Squared	-0.7	-0.4	-0.6	-0.3
	(0.1)***	(0.0)***	(0.1)***	(0.1)***
Some Secondary	81.2	-0.4	97.7	13.3
	(20.2)***	(12.9)	(27.5)***	(17.6)
Grade 11-13	72.4	70.4	13.1	2.6
	(19.2)***	(10.9)***	(28.3)	(24.4)
Some Post-Secondary	-14.8	14.9	-101.3	-31.9
	(26.6)	(15.4)	(39.4)**	(27.4)
Certificate or Diploma	64.7	49.4	56.3	-1.8
	(25.0)***	(15.6)***	(33.6)*	(27.8)
Bachelor/Master/PhD	118.2	154.6	7.3	55.5
	(21.4)***	(16.8)***	(33.3)	(35.0)
Foreign Born	-60.8	-71.4	-28.1	-3.2
	(19.5)***	(14.0)***	(26.0)	(28.3)
Missing Place of Birth	-311.4	192.4	-376.1	246.0
	(57.2)***	(83.8)**	(112.8)***	(130.2)*
Any Training	172.8	158.9	1,086.5	909.5
	(15.8)***	(11.4)***	(188.4)***	(205.3)***
Observations	10,469	11,938	10,469	11,938
R-squared	0.16	0.23		

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The estimated constant term is omitted.

Table A5
Complete Estimates for Bivariate Normal Model in Table 9
Estimated Coefficients (x100) and Probit Marginal Effects (x100)

	Weekly Earnings ($)		Training Incidence	
	Men	Women	Men	Women
Newfoundland	-168.9	-104.1	-0.40	-0.37
	(25.2)***	(12.3)***	(0.09)***	(0.09)***
PEI	-166.1	-36.6	-0.25	-0.24
	(25.0)***	(15.7)**	(0.10)**	(0.09)***
Nova Scotia	-125.8	-68.9	0.01	-0.10
	(20.5)***	(12.9)***	(0.07)	(0.07)
New Brunswick	-71.9	-50.4	-0.25	-0.21
	(21.3)***	(13.1)***	(0.07)***	(0.07)***
Quebec	-38.8	-46.6	-0.40	-0.50
	(21.1)*	(12.1)***	(0.06)***	(0.06)***
Manitoba	-40.8	-58.3	-0.07	0.00
	(20.3)**	(12.2)***	(0.07)	(0.07)
Saskatchewan	-93.9	-51.6	-0.05	0.02
	(20.3)***	(12.4)***	(0.07)	(0.06)
Alberta	-23.3	-35.5	-0.13	-0.01
	(21.7)	(12.9)***	(0.06)**	(0.06)
B.C.	-54.9	-40.9	-0.01	0.01
	(21.6)**	(15.2)***	(0.07)	(0.06)
Census Metro Area	43.6	28.5	0.04	-0.02
	(16.7)***	(10.6)***	(0.05)	(0.05)
Toronto/Montreal/Vancouver	11.5	60.2	-0.16	0.05
	(20.4)	(12.8)***	(0.06)***	(0.06)
Urban Centres	-9.8	11.0	0.06	0.07
	(21.9)	(13.5)	(0.07)	(0.07)
Rural Area	-93.3	-48.0	0.03	-0.02
	(16.7)***	(10.5)***	(0.06)	(0.06)
Age	60.3	33.6	0.01	0.07
	(4.8)***	(3.2)***	(0.02)	(0.02)***
Age Squared	-0.7	-0.4	0.00	0.00
	(0.1)***	(0.0)***	(0.00)	(0.00)***
Some Secondary	85.5	0.5	-0.05	0.14
	(20.9)***	(12.9)	(0.12)	(0.14)
Grade 11-13	57.2	66.3	0.25	0.46
	(20.7)***	(11.0)***	(0.08)***	(0.07)***
Some Post-Secondary	-37.0	12.0	0.29	0.20
	(28.2)	(15.5)	(0.08)***	(0.08)**
Certificate or Diploma	62.5	46.3	0.04	0.21
	(25.5)**	(15.7)***	(0.08)	(0.07)***
Bachelor/Master/PhD	89.8	148.6	0.27	0.32
	(23.8)***	(17.1)***	(0.06)***	(0.06)***

Table A5 (continued)
Complete Estimates for Bivariate Normal Model in Table 9
Estimated Coefficients (x100) and Probit Marginal Effects (x100)

	Weekly Earnings ($)		Training Incidence (%)	
	Men	Women	Men	Women
Foreign Born	-52.4	-67.3	-0.12	-0.30
	(20.1)***	(14.1)***	(0.06)**	(0.06)***
Missing Place of Birth	-328.0	195.6	0.22	-0.23
	(63.2)***	(85.1)**	(0.29)	(0.35)
Spouse Present			0.43	0.08
			(0.08)***	(0.08)
Spouse - Some Secondary			0.13	0.20
			(0.17)	(0.12)*
Spouse - 11-13			0.17	-0.01
			(0.08)**	(0.08)
Spouse - Some Post-Secondary			0.11	0.03
			(0.09)	(0.10)
Spouse - Certificate or Diploma			-0.09	0.05
			(0.09)	(0.10)
Spouse - Bachelors or More			0.20	0.06
			(0.07)***	(0.07)
1 Child			0.04	0.00
			(0.15)	(0.10)
2 or More Children			0.33	-0.04
			(0.22)	(0.11)
1 Child with Spouse			0.09	-0.08
			(0.16)	(0.12)
2 or More Children with Spouse			-0.27	0.03
			(0.23)	(0.13)
1 Preschool Child			-0.33	-0.18
			(0.29)	(0.12)
2 or More Preschool Children			-0.17	-0.25
			(0.10)*	(0.09)***
Preschool Children with Spouse			0.16	0.10
			(0.30)	(0.13)
Any Training	407.2	204.4		
	(90.4)***	(18.5)***		
Observations	10,469	11,938		
Log-Likelihood	-61962067	-58169147		
Rho	-0.32	-0.09		
Pr(Rho=0)	0.0088	0.0009		

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The estimated constant term is omitted.

Table A6

Complete Estimates for Linear Regression and IV Models in Tables 10A and 10B
Estimated Coefficients (x100)

	Linear Model		Instrumental Variable	
	Men	Women	Men	Women
Newfoundland	-186.8	-102.3	-112.0	-10.1
	(22.2)***	(12.4)***	(38.0)***	(38.7)
PEI	-171.6	-38.8	-86.0	21.3
	(22.5)***	(15.6)**	(41.6)**	(32.4)
Nova Scotia	-127.6	-69.7	-114.2	-16.5
	(19.5)***	(12.6)***	(39.0)***	(29.3)
New Brunswick	-84.1	-50.4	-33.3	23.4
	(19.5)***	(12.8)***	(34.0)	(28.3)
Quebec	-62.0	-49.5	35.0	86.3
	(18.0)***	(11.9)***	(42.3)	(37.0)**
Manitoba	-44.3	-53.9	1.4	-46.3
	(19.2)**	(11.8)***	(41.5)	(28.1)*
Saskatchewan	-102.0	-51.3	-100.4	-34.8
	(19.5)***	(12.1)***	(37.1)***	(25.3)
Alberta	-32.8	-35.5	-4.6	-21.8
	(20.4)	(12.7)***	(33.9)	(22.9)
B.C.	-49.6	-36.3	-6.0	-32.1
	(20.9)**	(14.8)**	(34.2)	(26.1)
Census Metro Area	50.2	28.0	51.8	36.4
	(15.9)***	(10.7)***	(24.2)**	(19.3)*
Toronto/Montreal/Vancouver	1.9	59.7	43.6	47.7
	(18.8)	(12.6)***	(30.0)	(20.9)**
Urban Centres	0.5	15.6	13.8	19.0
	(21.0)	(13.4)	(32.2)	(26.1)
Rural Area	-90.6	-49.8	-95.6	-32.9
	(16.3)***	(10.5)***	(24.3)***	(22.2)
Age	57.6	31.2	24.0	15.2
	(4.4)***	(3.1)***	(10.3)**	(10.9)
Age Squared	-0.7	-0.4	-0.3	-0.2
	(0.1)***	(0.0)***	(0.1)**	(0.1)
Some Secondary	82.1	3.4	85.3	28.5
	(20.4)***	(12.8)	(29.4)***	(19.7)
Grade 11-13	62.1	65.7	-7.7	-19.4
	(19.0)***	(10.6)***	(30.7)	(28.4)
Some Post-Secondary	-11.6	16.8	-41.6	-25.4
	(26.1)	(14.7)	(46.4)	(30.6)
Certificate or Diploma	62.3	48.7	38.8	-31.2
	(24.6)**	(14.9)***	(35.5)	(35.0)
Bachelor/Master/PhD	120.8	146.4	97.5	3.7
	(21.2)***	(16.6)***	(46.3)**	(43.0)
Foreign Born	-57.1	-60.5	-8.9	38.8
	(19.2)***	(13.6)***	(30.2)	(37.5)
Missing Place of Birth	-288.9	186.4	-367.6	233.6
	(56.1)***	(83.2)**	(135.0)***	(154.2)
Employer-Financed Training	256.6	256.2	1,162.5	1,112.1
	(17.1)***	(11.9)***	(203.6)***	(391.0)***
Government-Financed Training	-122.4	-93.0	720.1	-666.9
	(50.5)**	(23.3)***	(1,736.9)	(681.5)
Self-Financed Training	-60.5	15.1	-969.6	1,035.2
	(28.4)**	(19.2)	(646.0)	(679.1)
Observations	10,469	11,938	10,469	11,938
R-squared	0.18	0.27		

Notes:
Estimated (robust) standard errors appear in parentheses. The "*" denotes statistical significance at the ten percent level, "**" at the five percent level, and "***" at the one percent level. The estimated constant term is omitted.