

# The Knowledge Lift: The Swedish Adult Education Program that Aimed to Eliminate Low Worker Skill Levels

James Albrecht \*

Gerard J. van den Berg †

Susan Vroman ‡

February 5, 2005

## Abstract

The Swedish adult education program known as the Knowledge Lift is unprecedented in its size and scope, aiming to raise the skill level of all low-skilled workers towards the medium level. This paper evaluates the effects of program participation on individual labor market outcomes, notably employment and annual income, as well as on the labor market equilibrium. For the effects at the individual level, we apply fixed effect methods allowing for treatment effect heterogeneity. The data are based on a number of matched longitudinal administrative data sets covering the full population of Sweden. For the equilibrium effects, we analyze an equilibrium search model with heterogeneous worker skills. This model is calibrated using pre-program observations.

---

\*Department of Economics, Georgetown University; IZA; CESifo.

†Department of Economics, Free University Amsterdam; IFAU-Uppsala; Tinbergen Institute; CEPR; IZA; IFS.

‡Department of Economics, Georgetown University; IZA; CESifo

*Keywords:* returns to education, training, program evaluation, wages, participation, unemployment, schooling, Swedish labor market, selectivity bias, treatment effect.

*Acknowledgements:* We thank the Swedish National Labour Market Board (AMS) and Statistics Sweden (SCB) for their permission to use the data, and IFAU for financial support. We thank Maria Melkersson, Håkan Regnér, Per-Anders Edin, Katarina Richardson, Annette Bergemann, Bertil Holmlund, and seminar participants at IFAU/Uppsala University for helpful comments and suggestions. In addition, we thank Erika Ekström, Nadia Nilsson, and Helge Benmarker for research assistance and help with the data.

# 1 Introduction

Life-long learning, adult education, and employability have become focal points in the labor market policies of many advanced economies (see e.g. the recent OECD Employment Outlook 2004). It is expected that these economies will face more turbulent conditions than in the past, and that the development of novel production technologies will proceed at a sustained high speed. This would then require a flexible and suitably skilled workforce. Indeed, the role of low-educated workers has diminished in the modern knowledge-based economy. The fact that the age structure is changing towards a heavier representation of older workers means that the human capital adjustment needs to be made by the existing stock of workers instead of the inflow of new workers.

Sweden is relatively well prepared for a policy intervention to accommodate these adjustments, given its long tradition of training of adult unemployed workers (see e.g. Ministry of Education, 1998, Friberg, 2000, and Ministry of Industry, 2001). In 1997, Sweden implemented a new major adult education program called the “Adult Education Initiative” or “Knowledge Lift” (henceforth denoted as KL). Without exaggeration, this constitutes the largest and most ambitious skill-raising program ever. It aims to raise the skill level of all low-skilled workers to the medium skill level. It focuses on workers with a low level of education. The size of the program is unprecedented: in the period 1997-2000, more than 10% of the labor force has participated in it.

Obviously, the program reflects a great deal of optimism about the extent to which an adult’s human capital can be improved. The empirical literature on training programs for unemployed workers does not warrant this optimism. The general conclusion from this literature is that training does not have large effects on the individual’s labor market outcomes (see e.g. Fay, 1996, Heckman, LaLonde and Smith, 1999, and Martin and Grubb, 2001). A major exception concerns training for women who return to the labor market after a spell of child-raising activity, who clearly form an important target group for adult education. However, perhaps more importantly, training participants with a low initial level of education benefit even less than other educational groups.

In this paper, we analyze the effects of KL. The paper consists of two parts. In the first part, we estimate the effects on individual outcomes; in particular, on income and employment. In the second part, we calibrate the effects on labor market equilibrium using an equilibrium matching model with labor market frictions and skill heterogeneity.

For the empirical analyses, we use a rather unique set of longitudinal admin-

istrative data which contains the full population of individuals in Sweden. The dataset matches detailed records from employment offices, unemployment insurance agencies, the income tax agency, and the adult education register. The latter contains records of all adult education courses that are followed.

Longitudinal data on individual training program participation often display a pre-program earnings (or Ashenfelter’s) dip (see e.g. Heckman and Smith, 1999, for an exposition). Our data display prolonged pre- and post-program dips in earnings as well as employment rates. It is therefore essential to use data with a sufficiently long time span (in our case, 1991–2000). We estimate average treatment effects on the treated, and we focus on two outcome variables: annual earnings and employment. We apply fixed effect estimation methods, specifically, conditional difference-in-differences and conditional probit.

Given the size of the program, equilibrium effects are likely to be substantial. In addition to the effect on the individuals in the program, other workers in the economy are likely to be impacted through changes in wages and/or through changes in unemployment and employment probabilities.<sup>1</sup> We analyze the equilibrium effects of the knowledge lift by calibrating an equilibrium model of the labor market. Obviously, such a model needs to incorporate skill heterogeneity and equilibrium unemployment. Also, firms should be able to choose their production technology in the face of the prevailing skill distribution. We use a model based on Albrecht and Vroman (2002), which is a very concise and amenable model satisfying the above requirements. It assumes two worker types, low skill and medium skill, with the number of workers of each type taken as exogenous in a given market. There are frictions in the process by which unemployed workers and vacancies contact one another, and the surplus generated by a worker/job match is divided using the Nash bargaining solution. The flow output of a match depends on the skill level of the worker as well as on the job type as decided by the firm when it created the vacancy.

We use macro data from 1996 to calibrate the pre-knowledge lift economy. This allows us to set values for the unobserved parameters that drive the theoretical model. We then address the question: “Suppose the knowledge lift were to change the skill distribution in the economy in a particular way. Using the calibrated economy as a base, what would the effects be?” Specifically, we impute the post-KL skill distribution into the model and solve for the new steady-state equilibrium. We derive the wages for low-skill and medium-skill workers as well as their unemployment rates and employment in low-skill and medium-skill jobs.

---

<sup>1</sup>See e.g. Lise, Seitz and Smith (2002) for a general discussion of equilibrium evaluation of policy programs.

We also derive the overall unemployment rate, labor market tightness (vacancies over unemployment), the proportions of low-skill and medium-skill vacancies, and the equilibrium effects on the treated.

Of course, we do not address all the potential effects of KL. For example, Björklund et al. (2004) show that KL lead to a large flow of teachers from regular secondary education to adult education, and they argue that KL therefore may have generated substantial negative external effects on the quality of regular education. Such effects increase the social costs of the program beyond the amounts given below. However, addressing them would be beyond the scope of the present paper. We also do not aim to address the use of adult education by young individuals who left the regular school system with low educational levels, as a short-cut towards regular university education (see e.g. Björklund et al., 2004, and Ekström, 2003, for discussions). For this reason we exclude individuals aged below 25.

To date, a few studies have examined the effects of adult education in Sweden on individual labor market outcomes. Ekström (2003) estimates the effect on annual income of following adult education in the (pre-KL) early 1990s, using difference-in-differences. She finds no positive income effects for either gender. Several studies have compared individual labor market outcomes between unemployed individuals who enroll in KL and unemployed individuals who enroll in labor market training, using propensity score matching or IV methods (see e.g. Axelsson and Westerlund, 1999, and Stenberg, 2003). The results depend strongly on the outcome measure, the evaluation method, and the type of labor market training and subpopulation considered.

For the US, Jacobson, LaLonde and Sullivan (2003) estimate the effects of adult education using a sample of displaced prime-aged workers. They find sizable returns. However, as Björklund et al. (2004) argue, generalizing from evidence on US adult education programs is difficult because there are so many low-skilled individuals in the US, many of whom may have had insufficient human capital investment opportunities earlier in life. Indeed, the skill distribution in Sweden is more compressed than in the US (see Björklund et al., 2004, for an exposition).

The paper is organized as follows. Sections 2 and 3 describe the KL program and the data, respectively. In Section 4 we describe the empirical approaches that we employ to study the effects on individual outcomes, and we present the estimation results. We also report the sensitivity of the results with respect to a number of assumptions concerning the construction of the variables of interest. Section 5 contains the results of the equilibrium analysis, and Section 6 concludes.

## 2 The Knowledge Lift

As explained below, KL was run through the existing municipal adult education system (KOMVUX) and can be seen as a major qualitative and quantitative upscaling of KOMVUX. By now, many studies provide detailed descriptions of KL and/or KOMVUX and their participants. See, for example, The National Agency for Education (1999), Axelsson and Westerlund (1999), Skolverket (2001), and Stenberg (2003) for information on KL, and Skolverket (2001) and Ekström (2003) for information on KOMVUX. We therefore restrict ourselves here to a brief summary.

KL is by far the largest adult education program ever in Sweden. It ran from July 1, 1997 to December 31, 2002 (and was followed by a permanent expansion of the number of seats in KOMVUX). The objective was to increase the skill level of adult low-skilled workers to the medium skill level, thereby helping these individuals strengthen their position in the labor market. Here, low skilled means having an educational attainment below the level of a 3-year “gymnasium” degree, while medium skilled means having attained this level but not any levels beyond that. The 3-year gymnasium degree roughly corresponds to the upper secondary education level or senior high school. Since 1995 this is the lowest possible upper secondary school diploma, whereas before that many individuals left high school with a 2-year degree. The program particularly targets unemployed adult low-skilled workers. In fact, low-skilled employed workers and medium-skilled unemployed workers are also often eligible for KL, and the enrollees contain many low-skilled employed workers, working part time or full time.

Like KOMVUX, KL focuses on the enhancement of general skills (for example, English, Swedish, and mathematics), as opposed to specific skills needed for particular professions. However, part of KL can be spent in vocational courses and work placement. In principle, it is possible to combine upper secondary courses with studies at an elementary level or with a program organized by the National Labor Market Board for the unemployed. The curricula and grade criteria for the attainment of the medium skill level are roughly the same as in the regular upper secondary education system.

KL is organized at the municipal level, and is run through the KOMVUX system. The organization may be joint with other municipalities. A municipality may purchase the services of education providers and/or cooperate with them. However, the municipalities are responsible for admission into KL. A single course typically starts twice a year and covers a half-year term.

At the level of the individual, admission into KL and KOMVUX is in principle

free. The underlying view is that KL participation must be led by the demand for education. A participant should have ample scope for personal choice regarding the type of study and its timing and location. Whether one can participate in a desired course only depends on the availability of courses and on the entry skill level requirement. Recruitment of participants is sometimes carried out in cooperation with trade union organizations or local employment offices.

KL and KOMVUX participants may be eligible for a range of income grants and financial study support measures. Some enrollees receive “special education support” (UBS). The amount of financial support is equivalent to unemployment insurance (UI). UBS is only given to KL participants who are entitled to UI payments at date of entry into the program. Moreover, the worker must be between 25–55 years old inclusive at date of entry into the program and must study at the elementary or upper secondary level. The grant is typically given for a maximum of one year. Sometimes, special adult study assistance and funding are available as a combination of a grant and a loan. Many participants rely on other financial resources. An individual who is full-time in KL/KOMVUX is considered to be out of the labor force unless he/she earns income on the side.

The state channels funds to the municipalities to finance KL and KOMVUX. The amount of funding depends on the municipality’s unemployment rate and skill level distribution, and on the scope of the municipality’s program. A conservative estimate is that, in the first years of its existence, the state spent at least SEK 3.5 billion (US \$350M) per year on KL. This equals almost SEK 1000 per labor force participant in Sweden. The spending covered the creation of some 100,000 annual study slots. In practice, the funding was more than sufficient to meet the demand for KL (see Statskontoret, 1999). This fact is important for our analyses because it implies that there was no quantity rationing.

The following gives an indication of the size of the program in terms of numbers of enrollees. In the fall of 1997, 538,004 individuals (out of a population of 8M) were (i) aged between 25 and 55, and (ii) participated in the municipal adult education, or were unemployed (in the sense of actively searching), or participated in one or more training programs. About 220,000 of these participated in KL and/or KOMVUX, and of these about 56,000 received UBS. About 35,000 KL/KOMVUX participants were registered as unemployed, and another 5,000 participated both in KL/KOMVUX and in employment training. The number of registered unemployed, including those participating in KOMVUX and/or training programs was about 330,000. For comparison, the number of pupils in regular upper secondary school was about 300,000, while the number of individuals participating in employment training programs was about 40,000. The figures do

not sum to the total of 538,004 because some individuals fall into more than one category. Typically, the number of individuals enrolled in KL/KOMVUX is about 50% larger than the full-time equivalent of the number of occupied slots. This indicates that many enrollees are part-time participants. Skolverket (2001) provides a wealth of additional information on the composition of participants and courses.

Due to KL, the number of individuals enrolled in adult education became dramatically larger than in earlier years (the increase in the stock of participants was about 80%). The KL initiative also involved the improvement and modernization of teaching methodologies and pedagogy. For all practical purposes, KL and KOMVUX were indistinguishable in the period in which KL ran. Therefore, in the remainder of the paper, we often simply refer to KOMVUX as the program we evaluate, but it should be kept in mind that we evaluate it over the KL period, and we emphasize that the existing KOMVUX program includes courses that do not aim at the attainment of a medium skill level but rather an improvement within the class of low skill sublevels.

## **3 The micro data**

### **3.1 The data registers**

Our dataset is built on a random sample of 200,000 individuals from the population of inhabitants aged between 16 and 65 in Sweden on December 31, 1997. These individuals have been longitudinally traced in four different administrative registers covering (subsets of) the period 1990–2000. Our dataset matches the records of individuals across these registers.

The first register is the official Swedish register of income and wealth, called RAMS. It is obtained from yearly income tax declarations for the years 1990–2000. It includes information from the population register, which is used to create the sample. The register provides observations on an annual basis of various types of income that each individual may receive. Specifically, we observe individual wage incomes, incomes in the form of government subsidies (including UBS), and income from self-employment. The RAMS data also provide information on some individual characteristics.

The second register (AKSTAT) is from the unemployment insurance fund. It provides week by week information on the amount of unemployment compensation that is received. Together, AKSTAT and RAMS enable the construction of our measure of annual earnings or income for the years 1990–2000. This equals

the sum of all above-mentioned incomes in RAMS plus unemployment benefit income (all before income taxes). However, the AKSTAT data from before 1994 are not completely reliable.

The third register is the so-called HÄNDEL dataset, which is based on registers at the employment office and is compiled by the Swedish Labor Market Board. It includes all individuals who ever registered as unemployed starting from September 1991. Registration is voluntary but is required in order to receive or apply for unemployment compensation or to participate in any type of labor market program, so in fact almost all unemployed are in these data (according to Carling, Holmlund and Vejsiu, 2001, more than 90% of the individuals who are ILO-unemployed according to labor force surveys also register at the employment offices).

The HÄNDEL data provide labor market histories for all its individuals on a daily basis, with dates of transitions between different labor market states and between open unemployment and participation in training programs and work experience programs. However, because participation in KL is regarded as an out-of-the-labor-force activity, HÄNDEL by itself does not allow for observation of spells of KL participation. The HÄNDEL data also provide individual characteristics.

The fourth register (KOMVUX) contains individual records on participation in any adult education program. These data are available for the years 1990–2000. From this we can follow participation in adult education on a basis of six-month periods at the individual level. Therefore, for all individuals for all semesters there is a specific variable which says whether someone has been in KOMVUX in that semester. This includes those whose participation is subsidized as well as those who do not get any subsidy. There is also information about whether the course(s) taken were day or evening courses, about the level of education prior to participation, and about the municipality where the course was taken. For the years 1997–2000, additional detailed information on adult education experiences is available, but this information is incomplete and could not be satisfactorily matched to the KOMVUX register, with one important exception: we observe for each course taken whether it was completed or whether the participant dropped out.



## 3.2 Variable definitions, sample selection, and data descriptives

We focus on the treatment of being in KOMVUX in the KL years. In particular, we distinguish between the following treatments: being in KOMVUX in the second half of 1997 (97-II) and not before or after, being in KOMVUX in 98-I and not before or after, being in KOMVUX in 98-II and not before or after, as well as two-semester treatments, 97-II and 98-I and not before or after, 98-I and 98-II and not before or after. We do not consider treatments of more than two semesters, because these often involve different remuneration eligibilities. We also do not consider treatments consisting of KOMVUX participation periods that are interrupted by semesters out of KOMVUX, since such treatments may have fundamentally different effects. Also, we do not consider treatments that end after 98-II because then the time distance to the (only available) post-treatment year 2000 is too small to detect any long-run effect. Finally, we do not consider treatments where no course is completed by the end of the semester in which it should have been completed. The individuals with such incomplete treatments are dropped from the data.

We restrict attention to individuals who are between 25 and 55 at the moment of treatment. This is because individuals below 25 and above 55 face different active labor market programs, educational opportunities, and remuneration eligibilities while unemployed and in education (see e.g. Larsson, 2003). As noted in Section 1, young individuals who left the regular school system with low educational levels may use KOMVUX as a short-cut towards regular university education (see e.g. Ekström, 2003). In such cases, it takes many years after participation in KOMVUX before the individual returns to the labor market.

We also restrict attention to low-skilled individuals, i.e. having at most two years of upper secondary education or the equivalent at the start of the treatment. This restriction is consistent with the main KL objective to raise the skill level of low-skilled workers. For each treatment considered, we exclude individuals who were in KOMVUX at any time between 1990 and the beginning of the treatment, or who were in KOMVUX at any time after the treatment up to the end of the observation window. However, we make an exception to this rule if the data unambiguously establish that the KOMVUX participation concerned the attainment of a level lower than or equal to two years of upper secondary education, as this constitutes a transition within the low-skill category.

The data contain only a few explanatory variables at the individual level. Most of the analyses are carried out with data stratified on gender and (at the

same time) on whether one is prime-aged (25–40) or older (41–55), with age measured at the year of the treatment considered. Even then, the number of individuals by gender  $\times$  age category who are treated in certain semesters is rather small (see below). We therefore adopt empirical approaches that do not require many explanatory variables (see the next section). We discard the group of older men because the number of treated observations is too small for any meaningful analysis. For the same reason we discard immigrants, except those from Nordic countries, who are merged with the natives (about 18% of the participants are immigrants from outside the Nordic region.)

For a given treatment period and age interval and gender, the control group is defined by the restriction that its individuals satisfy the same criteria as the treated (being of the same gender, being in the same age interval in the treatment period, having a low skill level, satisfying the KOMVUX participation restrictions, and having been born in Sweden or another Nordic country) except of course that they should not be in KOMVUX during the treatment period. Note that one individual can be a member of control groups for different treatment periods. Members of the treatment and control groups can work during the treatment periods.

To capture the pre-treatment conditions of individuals, we focus on outcomes in 1994. This year is sufficiently far before the treatment periods to rule out anticipatory effects. For the years before 1994 the unemployment insurance data are less reliable, and for the years before 1993 the employment office data are relatively unreliable. The year 1996 displays the pre-program earnings (or Ashenfelter) dip. Figure 1 illustrates this concerning being in KOMVUX in 97-II by way of a graph of the annual income of individuals between 25 and 40 for the treated and the controls (in the price level of the year 2000). The figure also suggests a post-treatment earnings dip in at least 1998 and (for men) 1999. For the longer treatment periods, the post-program earnings dip seems to be more pronounced and prolonged (see e.g. Figure 2 concerning being in KOMVUX in both 97-II and 98-I).<sup>2</sup> We therefore primarily use 2000 as the post-treatment evaluation year.

Table 1 shows the effect on the sample size of successive imposition of the data restrictions discussed above, when we apply successive restrictions to the treatment and control group samples for treatment 97-II using 1994 and 2000 as pre- and post-program years. In this case, we end up with treatment and control groups of 225 and 38,670 individuals, respectively. Subsequent stratification into

---

<sup>2</sup>For young individuals such a dip may result from using KOMVUX as a bridge to subsequent regular university study for a number of years. However, recall that in the data we exclude individuals aged below 25.

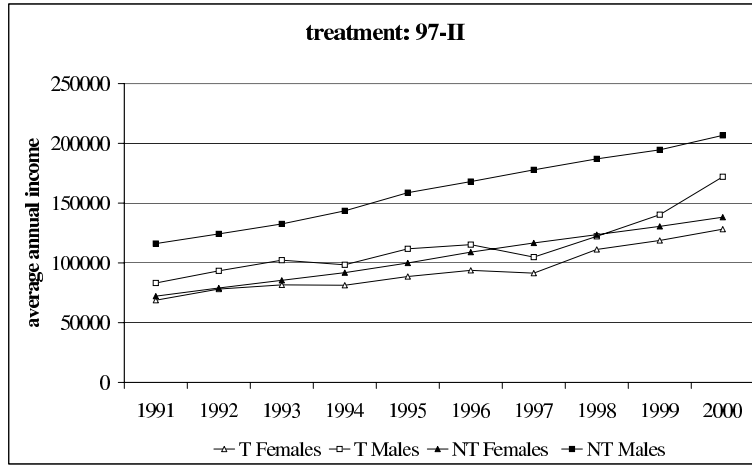


Figure 1: Average annual income for treated and controls aged 25–40, concerning treatment in 97-II.

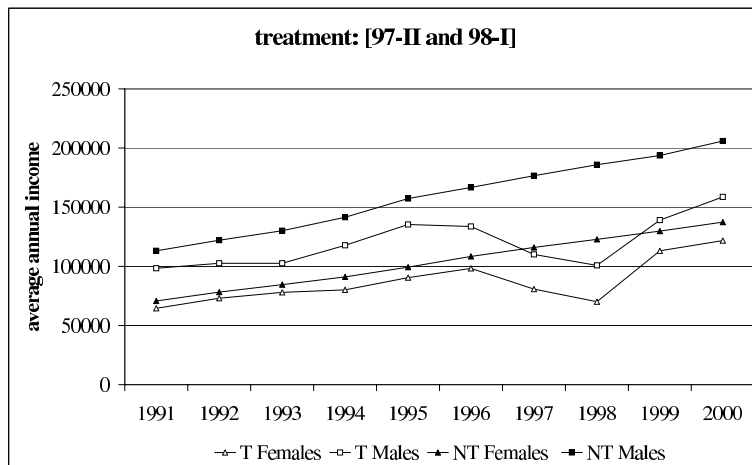


Figure 2: Average annual income for treated and controls aged 25–40, concerning treatment in [97-II and 98-I].

Table 1: *Determination of the ultimate samples for the evaluation of participation in KOMVUX in 97-II, using 1994 and 2000 as pre- and post-program years.*

successive selection criterium	resulting sample size
from population aged between 16 and 65 in 1997	200,000
present in this age frame in 1994 and 2000	175,221
aged between 25 and 55 in August 1997	131,352
born in Sweden or other Nordic countries	120,060
low skilled in August 1997	69,414
no KOMVUX before or after 97-II	55,295
of which:	
→ in KOMVUX in 97-II	333
no dropout	254
not men aged 41-55 ( $\Rightarrow$ <i>treatment groups</i> )	225
→ not in KOMVUX in 97-II	54,962
not men aged 41-55 ( $\Rightarrow$ <i>control groups</i> )	38,670

gender  $\times$  age categories results in subsamples with treatment groups of around 70 individuals.

For most treatment periods we consider, the average income level in 1994 among treated men is lower than the average income level in 1994 among the corresponding controls. In most cases the difference is significant. However, among young and older women the difference is neither systematically positive nor negative, and the differences are not significant. The same applies to the employment rates and to average outcomes in 2000. Thus, earnings and employment probabilities among the treated men are on average lower than in the corresponding population of low-skilled men. Clearly, therefore, the treated men are a selected subgroup of this population. For women we do not find such evidence.

The years before 1994 in the annual-income observation window (1990–1993) coincide with a major recession in Sweden.<sup>3</sup> Note from Figures 1 and 2 that the

---

<sup>3</sup>In 1994 the unemployment rate was still relatively high. Our estimation methods take

average income among treated men decreases at the end of this recession and does not recover in the years immediately after that. This also shows up for the other treatment periods we consider and for the male employment rates. Apparently, the set of males who enrolled in KL in the late 1990s contains relatively many men who suffered from the recession in the early 1990s in the sense that their labor market position was persistently worse afterwards. This fits in with the “life-long learning” objective of modern adult education programs to help individuals who are affected by negative shocks. For young women, the drop in average income at the end of the recession is smaller, and for older women we do not find any evidence of a drop. This again suggests that (older) female enrollees have a different background than the male enrollees.

## 4 Empirical analysis of average effects on individual outcomes using micro data

### 4.1 Effects on income

#### 4.1.1 Conditional difference-in-differences estimation of average treatment effects on the treated

To analyze the effects on annual incomes, we apply a version of the “conditional difference-in-differences” method developed by Heckman et. al. (1998)<sup>4</sup>. This method estimates the average treatment effect on the treated (ATET) in a population, allowing for treatment-effect heterogeneity and self selection on unobservables. In order to estimate the average effect, the individuals in the treatment and control groups should be aligned in terms of observed explanatory variables, for example, by matching them on the propensity score. The average difference between the treatment and control groups of the differences in the outcomes in the pre- and post-treatment evaluation years then equals the ATET. To see the assumptions involved, it is useful to consider briefly a regression specification for annual income  $Y_{i,t}$  of individual  $i$  at year  $t$ . This is conceptually inferior to a nonparametric counterfactual framework (see e.g. Heckman, LaLonde and Smith, 1999) but it facilitates the exposition. Let

$$Y_{i,t} = g(X_i, t) + V_i + \delta(U_i, V_i, X_i)I(Z_{i,t} > 0) + \varepsilon_{i,t} \quad (1)$$

---

account of changes in aggregate conditions.

<sup>4</sup>See Bergemann, Fitzenberger and Speckesser (2004) and Blundell et al. (2004) for applications.

where  $X_i$  are observed explanatory variables,  $g$  is an unknown function,  $U_i$  and  $V_i$  are individual-specific effects (which may depend on  $X_i$ ),  $\delta$  is the treatment effect function (which may depend on  $t$ ),  $Z_{i,t} > 0$  indicates that the individual has participated in KL in the past, and the errors  $\varepsilon_{i,t}$  are zero-mean and independently and identically distributed. Note that we may allow for time-dependent observed explanatory variables, and we may allow  $\delta$  to depend on other covariates than those in  $g$ . For each  $i, t$  we observe  $Y_{i,t}, X_i$ , and  $I(Z_{i,t} > 0)$ . Participation in KL is endogenous in the sense that it may be affected by the same individual characteristics as those affecting  $Y_{i,t}$ . Specifically,  $Z_{i,t}$  may be affected by  $U_i, V_i$ , and  $X_i$ , but it is independent of  $\varepsilon_{i,t}$ . In the regression specification, the spurious dependence between treatment participation and outcome runs by way of  $V_i$  and the spurious dependence between treatment effect and outcome runs by way of both  $U_i$  and  $V_i$ .

Let the pre- and post-treatment evaluation years be denoted by  $t = 0$  and  $t = 1$ , respectively. Consider two individuals ( $i = a, b$ ) with the same  $X$ , but  $Z_{b,1} > 0$  whereas  $Z_{a,1} < 0$ . From (1) it follows that

$$(Y_{b,1} - Y_{b,0}) - (Y_{a,1} - Y_{a,0}) = \delta(U_b, V_b, X_b) + \varepsilon_{b,1} - \varepsilon_{b,0} - \varepsilon_{a,1} + \varepsilon_{a,0}$$

so that, by taking the average over the treated, we obtain the ATET. Clearly, it is crucial that the income variable is additive in the unobserved fixed effects  $V_i$ . Moreover, these fixed effects should not change over time. The average time trend effects in income should be the same for individuals in the treatment and control groups who have identical covariates.

In our version of conditional difference-in-differences estimation, the analysis is stratified on all observed explanatory variables. In a given analysis, the treated and controls are then homogeneous in terms of observed explanatory variables (but not in terms of unobserved explanatory variables). This means that the ATET simply equals the difference-in-differences of the means of  $Y$  calculated for the treated and controls for the pre- and post-treatment year.

Conditional difference-in-differences has a number of advantages. First, it allows for effect heterogeneity. Second, it can handle Ashenfelter's Dip in a natural way. Third, it handles various types of selection effects. This includes self-selection into the treatment on the basis of the expected individual outcomes and costs, and selection of the moment of enrollment and the duration of participation. Fourth, it does not require exclusion restrictions (as in natural experiments) to identify the treatment effect. Such a restriction requires that the data contain a variable that affects the treatment assignment but that does not affect the outcome of interest other than by way of the treatment. This is often difficult to

justify. If a variable is observed by the analyst, then it is often also observable to the individuals under consideration. If the variable affects the probability of treatment, and the individual knows that he may be subject to treatment, then he takes his value of the variable into account to determine his optimal strategy, and this strategy in turn affects the rate at which the individual leaves the state of interest. Our data certainly do not contain any candidate instrumental variables.

Our approach of stratifying the treatment and control groups by age and gender is motivated by a paucity of observed explanatory variables. However, it also has some specific advantages over the use of propensity score matching. First, the results do not depend on a functional form for the propensity score as a function of the matching variables. Second, we do not need to rely on asymptotic theory or bootstrapping in order to estimate standard errors. Instead, we may calculate the exact small-sample standard errors of the estimated effects.

The assumption that the average time trend effect in income does not depend on the treatment status of individuals with identical covariates may be violated in practice. On average, non-treated individuals gather more work experience than treated individuals between the pre- and post-program year, for three reasons. First, they have more time to work during the treatment period. Second, they are less often unemployed shortly before the treatment period. Third, many of them work at the end of the treatment period, so they do not need to spend time to find a job then. With positive returns to work experience, their long-run average income levels may exceed those of initially identical individuals who were treated, even if the treatment itself does not have any causal effect. In the presence of a positive treatment effect, this would lead to an underestimate of this effect. This is a common but ignored problem in applications of (conditional) difference-in-differences.

#### **4.1.2 Estimation results**

Table 2 presents the conditional difference-in-differences estimates of the average treatment effect on the treated concerning participation in KOMVUX in 97-II, using 1994 and 2000 as pre- and post-program years. Clearly, the ATET estimates are all insignificantly different from zero (note that the estimates for the young groups are in accordance with Figure 1). Concerning the magnitude of the estimates, it should be recalled that young men on average have higher wages than women and that the outcome variable of interest is the level rather than the logarithm of income.

We also find insignificant ATET estimates for other treatments with a dura-

Table 2: *Estimates of the average treatment effects on the treated, concerning participation in KOMVUX in 97-II, using 1994 and 2000 as pre- and post-program years and annual income as outcome measure (in 1000 SEK in the year 2000).*

type of individual	ATET	(standard error)	$n_{treated}$	$n_{control}$
Male, aged 25–40	10.5	(13)	69	15061
Female, aged 25–40	0.6	(10)	96	9163
Female, aged 41–55	-3.6	(13)	60	14446

tion of one semester. Note that for those cases the time between the treatment and the post-program year is at most only 1.5 years. From the previous section, comparing incomes in 2000 for such cases is hampered by the post-treatment earnings dip.

For the treatments with a duration of 1 or 1.5 years, the above conclusions are reinforced. The results are also insensitive to whether medium-skilled individuals are included among the treated. We repeated the analysis using log income instead of the income level as outcome variable, but this did not affect the conclusions.

We also performed analyses with trimmed income data. This serves to exclude any disturbance effects of measurement errors in the data. Moreover, very low incomes – including zero incomes – do not reflect the individual monetary means available for survival; presumably such individuals rely on family members’ income. We use a lower income bound of SEK 50,000 per year for every year, and an upper bound equal to SEK 160,000 in 1994, linearly increasing with SEK 10,000 per year, up to SEK 220,000 in 2000 (all these numbers are in terms of the 2000 price level). These bounds are applied to the treated as well as the corresponding controls. We may therefore also consider the trimming as a crude device to match treated and controls on the pre-treatment annual income.

Table 6 in Appendix 1 gives the ATET estimates for the individuals whose incomes in the pre- and post-treatment years were within the above range. Again, all estimates are insignificant. In comparison to the estimates in Table 2, the effects are somewhat larger for women and smaller for young men. This suggests that men without income (and men with very low incomes) benefit more from KL than men with better labor market conditions. For women, the reverse holds.



We also examine the effects on the median of the income distribution. For each gender, age, year, and treatment semester(s), one may replicate the ATET estimation with median incomes instead of average incomes. The main advantage is that this is insensitive to the handling of zero incomes and income outliers. The main disadvantage is that the resulting estimate cannot be related to a meaningful effect at the individual level, basically because the difference of the median does not equal the median of the difference. The estimates on the effect on the median somewhat exceed those in Table 2.

The results suggest that KL does not increase annual income. Alternatively, the sample sizes of treated individuals may simply be too small for a meaningful analysis. There is some evidence that disadvantaged young men benefit from KL in terms of annual income.

### 4.1.3 Meta analysis

Thus far we have performed separate analyses for different treatment periods and for different gender  $\times$  age types. However, the time dependence of the average income of the various treatment and control groups satisfies certain regularities (recall the figures in the previous section). We may exploit this to enhance the efficiency of the analysis. After all, the insignificance of the non-parametric estimates might be due to the small sample sizes for given treatment periods. In addition, recall that treatments in periods that end close to the latest year available in the data are hard to evaluate, but if the pre-treatment-period income patterns are similar to those for other treatment periods then we can still use the corresponding data to enhance efficiency.

In this subsection, we therefore analyze the treatment effects by gender  $\times$  age type and treatment period in a unified framework. We postulate a parametric non-linear regression model for the observed average annual income  $Y$  per gender  $\times$  age type, per year, per treatment period, and per treatment status, where the average is taken over the individuals in the specific group we consider. This model is supposed to capture the average income patterns over time for the various treatment and control groups. It therefore also captures the ATET's that we estimated nonparametrically in the previous subsection.

We now let the index  $i$  denote the gender  $\times$  age type instead of the individual, and we again use  $t$  to denote a given year, the value  $t = 1$  corresponding to the year 1991 and the value  $t = 10$  to 2000. We distinguish the various treatment periods by way of variables  $a$  and  $\tau$ . Here,  $a$  denotes the length of the treatment period ( $a = 0$  and  $a = 1$  corresponding to lengths 0.5 and 1, respectively) and  $\tau$  denotes the center of the period, as measured by the calendar year in which the

income dip is deepest (1998, except for the treatment period 97-II). Finally, we use  $D$  to denote the treatment status of the group considered. This is a binary variable taking the value 1 if the individuals in the group will be treated in one of the treatment periods. We now postulate

$$\begin{aligned}
Y_{i,t,a,\tau,D} = & c_{0,i} + c_{1,i}t + c_{2,i}D + & (2) \\
& [c_3a + c_4(1 - a)] \cdot D \cdot e^{-\frac{|t-\tau|}{c_5}} + \\
& c_{6,i} \cdot D \cdot I(t \geq \tau) \cdot \left( \frac{e^{t-\tau} - 1}{e^{t-\tau} + e} \right) + \varepsilon_{i,t,a,\tau,D}
\end{aligned}$$

Consider first the control groups, i.e. with  $D = 0$ . The parameters  $c_{0,i}, c_{1,i}$  describe for each  $i$  separately a linear equation for  $Y$  as a function of  $t$ . The parameters  $c_{2,i}$  give the long-run pre-treatment difference between treatment and control groups for each  $i$ . The values of  $c_{2,i}$  reflect the extent to which the treated are a special subgroup of the individuals with a given  $i$ . Note that we assume that these parameters do not depend on the length or starting date of the treatment period.

The parameters  $c_3, c_4$  and  $c_5$  capture the pre- and post-program dip. If  $c_5$  is large then the dip is wide. The functional form of the shape of the dip as a function of  $t - \tau$  and  $a$  reflects the shape observed in the data. For example, if the treatment period is long, then the dip is deeper. The shape of the dip is constrained to be symmetric and is also assumed to be identical across  $i$ .

Finally, the parameters  $c_{6,i}$  capture the ATET's. The corresponding ratio in parentheses reflects the fact that the ATET is only revealed after the post-treatment dip. This term increases in  $t$ , from 0 at  $t = \tau$  to 0.31 at  $t = \tau + 1$ , and so on, quickly converging to 1. For sake of brevity we do not make the ATET's dependent on the length or starting date of the treatment period.

Note that  $a$  and  $\tau$  and the associated parameters  $c_3, c_4, c_5, c_{6,i}$  are only relevant if  $D = 1$ . In general, the parameters of the equation should not be given a causal or structural interpretation. They merely capture how  $Y$  varies with its determinants. The only exception are the  $c_{6,i}$  parameters, which are supposed to capture the ATET's.

We estimate the equation with the data on  $Y$  for each combination of 3 values of  $i$ , 10 values of  $t$  (1991–2000), 5 different treatment periods, and 2 different treatment statuses. In fact, we only use data from one control group for each value of  $i$  and  $t$ , because the different control groups contain mostly the same individuals (accordingly, the data on  $Y$  are virtually the same across these groups). We use

the control group corresponding to the treatment period 98-I. As a result, we have  $3 \times 10 \times (5 + 1) = 180$  observations for the regression.

For older women, the data show that the linear specification of  $Y$  as a function of  $t$  in the control group is incorrect because  $Y$  flattens out as time proceeds. We therefore replace it by an increasing concave function. Specifically, we replace  $c_{1,i}t$  in equation (2) for older women by

$$c_1 \cdot \log(t + c_7)$$

Note that this makes the  $c_0$  and  $c_1$  parameters incomparable between young men and women on the one hand and older women on the other.

We estimate the model with Nonlinear Least Squares. The regression analysis ignores a few statistical complications. First, for a given gender, age interval, treatment period, and treatment status, the average incomes in different years mostly concern the same individuals. At the individual level, incomes in consecutive years are typically correlated, leading to dependency of the error terms in the average income regression. Second, the standard error of the estimated average income depends on the sample size in the year considered and therefore varies across years, leading to heteroskedasticity.

Table 3 presents the estimation results. The main result concerns the insignificance of the ATET's for all age and gender types. This reinforces the results of the previous subsection. Other notable results are that young male enrollees are a selected subset of the corresponding population, whereas this is not true for young and older women. The pre- and post-program dip is deeper for longer treatments. The average annual population income increase is larger for young men than for young women, and this is in turn larger than for older women.

The results are very robust. Notably, the insignificance of the ATET's is robust with respect to changes in the functional form of the dip and the time dependence of the extent to which the ATET reveals itself. Also, if we restrict attention to the calendar years 1994–2000, then we obtain virtually the same results.

Differences in annual income are an imperfect indicator of productivity differences, because they also reflect differences in the employment rate. In the next subsection we directly estimate the employment effects of KL.

Table 3: *Estimates of the non-linear regression model for average annual income as a function of gender, age, calendar time, treatment period, and treatment status (in 1000 SEK in the year 2000).*

parameter	estimate	(standard error)
<i>constant term <math>c_0</math></i>		
all	36.6*	(11)
additional for male, aged 25–40	73.8*	(12)
additional for female, aged 25–40	25.2*	(12)
<i>time trend</i>		
$c_1$ male, aged 25–40:	8.8*	(0.6)
$c_1$ female, aged 25–40:	7.5*	(0.6)
$c_1$ female, aged 41–55:	43.1*	(4.3)
$c_7$ female, aged 41–55:	2.16*	(0.4)
<i>pre-program difference of treated <math>c_2</math></i>		
Male, aged 25–40	-27.7*	(3.3)
Female, aged 25–40	-4.3	(3.3)
Female, aged 41–55	0.0	(3.3)
<i>Ashenfelter dip</i>		
$c_3$ long treatment period	-50.1*	(4.0)
$c_4$ short treatment period	-19.8*	(3.5)
$c_5$ width of the dip	0.63*	(0.1)
<i>treatment effect <math>c_6</math></i>		
Male, aged 25–40	-6.0	(10)
Female, aged 25–40	-14.6	(10)
Female, aged 41–55	-1.4	(10)
R-squared	0.91	
# observations	180	

Note: \* denotes significance at the 5% level.

## 4.2 Effects on the employment rate

### 4.2.1 Conditional probit analyses of treatment effects on the employment rate

We distinguish between two labor market states: employment and unemployment. We are interested in the effect of the treatment on the probability of being employed. The model framework and estimation strategy are similar in spirit to those in the previous subsection. The main difference concerns the fact that now the outcome measure  $E$  is binary. This makes it less attractive to adopt a linear model for the outcomes, as, for example the linear model for  $Y$  in equation (1). We therefore specify a latent variable model and a corresponding analog of the conditional difference-in-differences estimation method.

Let the (latent) variable  $E_{i,t}^*$  satisfy, in obvious notation,

$$E_{i,t}^* = g(X_i, t) + V_i + \delta(X_i)I(Z_{i,t} > 0) + \varepsilon_{i,t} \quad (3)$$

The main restriction, as compared to the income effects analyses, is that the treatment effect is now constant for given  $X_i$ . That is, the treatment effect heterogeneity is reduced to heterogeneity across age and gender. This is the price to be paid for adopting a latent variable framework of analysis. For each  $i, t$  we observe  $E_{i,t} := I(E_{i,t}^* > 0)$  as well as  $X_i$  and  $I(Z_{i,t} > 0)$ . Participation in KL is endogenous in the sense that  $Z_{i,t}$  may be affected by  $V_i$  and  $X_i$ . Note that now the probability of employment is a non-linear function of the treatment status and the other determinants.

As in the previous subsection, we stratify the empirical analysis on all observed explanatory variables  $X_i$ . In a given analysis, the treated and controls are therefore homogeneous in terms of observed explanatory variables (but not in terms of unobserved explanatory variables). We may therefore simplify equation (3) to  $E_{i,t}^* = g(t) + V_i + \delta I(Z_{i,t} > 0) + \varepsilon_{i,t}$ .

The well-known conditional logit estimation method can be used to estimate the  $\delta$  parameter, in the following way. First, assume that the above outcome equation for given  $t$  defines a logit model. Secondly, let  $t = 0, 1$ , and consider the conditional probability of making a transition from  $E_{i,0} = 0$  to  $E_{i,1} = 1$  (or from  $E_{i,0} = 1$  to  $E_{i,1} = 0$ ) *given* that the individual labor market states at  $t = 0$  and  $t = 1$  are different. This defines a new binary outcome. We define a corresponding binary variable  $H_i$  as follows:  $H_i = 1$  if  $E_{i,0} = 0$  and  $E_{i,1} = 1$ , and  $H_i = 0$  if  $E_{i,0} = 1$  and  $E_{i,1} = 0$ . In all other cases,  $H_i$  is undefined. In words,  $H_i = 1$  if the individual makes a transition from unemployment to employment and  $H_i = 0$  if

the transition is the other way round, whereas individuals who are in the same state are removed. Note that what matters is not exactly what occurs between the years corresponding to  $t = 0$  and  $t = 1$ , but whether, and if so how, the individual labor market states differ between these years. The variable  $H_i$  also follows a logit model specification, with the treatment status as the only explanatory variable. We can thus introduce a latent variable  $H_i^*$  satisfying  $H_i^* = \beta_0 + \delta I(Z_{i,1} > 0) + \epsilon_i$ . This allows for straightforward estimation of  $\delta$ .

Note that this estimation method is similar to the conditional difference-in-differences method. Examining  $H_i$  instead of  $E_{i,t}$  means examining a sort of first difference across time for a given individual. It removes the fixed effect  $V_i$  from the analysis. The average of  $H_i$  in the control group provides an estimate of the time trend ( $\beta_0$ ) in the employment outcomes. The average of  $H_i$  in the treatment groups then identifies the treatment effect. In fact, we assume that  $H_i$  follows a probit model specification. This assumes a different class of distributions for  $\epsilon_{i,t}$  (see Magnac, 2004) but is otherwise equivalent to conditional logit.

A disadvantage of conditional logit and conditional probit is that they do not enable the estimation of the quantitative effect of treatment on the employment outcomes  $E_{i,t}$ . This effect is non-linear and depends on the unknown fixed effects  $V_i$ . This means that we can only make a qualitative evaluation, in the sense that we can only determine the sign and significance of the treatment effect.

In the empirical analysis, we divide the pre- and post-treatment year into three periods each: January-April, May-August, and September-December. We define an individual to be employed in a given period if he is employed for more than half of the period (otherwise we call him unemployed in the period). However, we do not directly observe the employment spells within a period. The fraction of time spent in employment has to be inferred from the observations of the moments of the transitions into and out of unemployment and the corresponding origin and destination states, notably employment, and from the income data. We cannot rule out the possibility that a period between successive transitions into and out of employment also includes spells of non-participation.

We compare the employment outcome in a given period in the pre-treatment year to the outcome in the corresponding period in the post-treatment year. This effectively multiplies the sample size by three (although only observations where the pre- and post-treatment year outcomes differ contribute to the empirical results). The three observations for a given individual may be dependent to the extent that their joint determinants are not captured by the common fixed effect. We conjecture that such a dependence would lead to an under-estimation of a positive employment effect, because if individuals make relatively many transi-

tions within a given year, then a transition from unemployment to employment for given periods of the year is often succeeded by a transition from employment to unemployment for the subsequent periods.

#### 4.2.2 Estimation results

Table 4 gives the results of the conditional probit analyses. “Fraction  $U \rightarrow E$  among treated = 0.74 ” means that among the treated who had a different individual labor market status in 1994 than in 2000, 74% were unemployed in 1994 and employed in 2000, and 26% were employed in 1994 and unemployed in 2000. For women aged below 40, the fraction of transitions from unemployment to employment is higher among the treated than among the controls, and this shows up in the estimate of  $\delta$  being positive.<sup>5</sup>

We find a significantly positive employment effect for young men. For young women the effect is positive but insignificant. For older women the effect is negative but very insignificant. Somewhat loosely, the result for young men is driven by the relatively high frequency of young men who are unemployed in 1994, follow KOMVUX in 97-II, and are employed in 2000.

We end this section by summarizing the conclusions of the analyses with micro data. First, KL has no significant effect on average income and employment of women. Among young and older women, enrollees are similar to non-enrollees in terms of average pre-program labor market outcomes, and KL does not improve the average outcomes of the enrollees. If anything, young women benefit slightly from KL, whereas KL participation of older women goes along with a slight deterioration of outcomes.

For young men the results are different. The enrollees are on average more disadvantaged in terms of pre-program labor market outcomes than the non-enrollees. In particular, they include relatively many who were hurt by the early-1990’s recession. On average, KL participation significantly increases the probability of employment among young men. We do not find a significant effect on average annual income, although the magnitude of the corresponding average treatment effect on the treated is much larger than for young and older women.

At first sight, this absence of a significant average income effect suggests that KL did not have positive productivity effects. This would be at odds with the objective of KL, which was to improve skill levels. A first explanation for the absence of an income effect in conjunction with the occurrence of an employment

---

<sup>5</sup>In a linear probability model, the estimated employment effect simply equals the difference of 0.74 and 0.68.

Table 4: *Estimates of the treatment effect on the employment probability, concerning participation in KOMVUX in 97-II, using 1994 and 2000 as pre- and post-program years.*

type of individual:	female, aged 25–40	female, aged 41–55	male, aged 25–40
fraction $U \rightarrow E$ among treated	0.74 (0.045)	0.39 (0.068)	0.90 (0.037)
number of transitions among treated	94	51	63
fraction $U \rightarrow E$ among controls	0.68 (0.0064)	0.42 (0.0062)	0.74 (0.0044)
number of transitions among controls	5347	6285	9977
probit estimate of $\beta_0$	0.47 (0.018)	-0.21 (0.016)	0.64 (0.013)
probit estimate of $\delta$	0.19 (0.14)	-0.064 (0.18)	0.67 (0.22)

effect is that the sample size on incomes among the treated is simply too small. A second explanation is that the only effective component of KL is the increased exposure to the labor market due to a strong involvement of the case worker who may bring the enrollee in contact with suitable vacancies. KL would then work as a job search assistance program (see Richardson and Van den Berg, 2001, for evidence that another Swedish training program works this way). However, if this is correct, then one would also expect an employment effect for female enrollees. Moreover, one would expect a strong employment effect very quickly after leaving the program.

A third explanation is that in Sweden, wages in the labor market do not reflect productivity, because of strong wage compression. In particular, the wage-enhancing effect of productivity increases due to KL for previously low-skilled workers may be constrained by Swedish labor market institutions governing wage



setting. However, this is hard to reconcile with the observed cross-sectional wage dispersion and the observed variation in individual wage changes over time. Also, with this explanation one would still expect an employment effect for female enrollees.

A fourth explanation is that the conditional difference-in-differences evaluation method may under-estimate the average KL income effect, because of the gain in work experience in the control groups. To get some feeling for the magnitude of the bias, suppose that, in 2000, the young males in the control groups have on average accumulated one additional year of regular work experience. According to Table 3, this corresponds to a gain of about SEK 8,800 in annual income in 2000. If we add this to the ATET estimate of 10.5 (i.e., 10,500 SEK) for young men taken from Table 2, then we obtain a corrected ATET estimate of 19.3 (standard error 13) in annual income (for women the corrected estimates are still close to zero). Any underestimation of income effects is exacerbated if the post-program year used in the evaluation is still in the post-program dip. In that case, the time frame of the data does not allow for the full effects of the program to be observed.

A fifth explanation is that there may be equilibrium effects that affect the average income effects. After all, the program is so large that a substantial fraction of the low-skilled workers without work participate in it. One may expect effects on the behavior and outcomes of employers and labor market participants who have not been enrolled in KL. This in turn may affect the outcomes of the enrollees as well. To investigate this, we perform an equilibrium analysis in the next section.

## 5 Analysis of equilibrium effects

### 5.1 The model

As indicated in the introduction, we use the equilibrium labor market model from Albrecht and Vroman (2002). We first present a generalization of that model. Then we calibrate it, and we simulate the potential equilibrium effects of the program.

The model is a stylized one in which risk-neutral workers live forever. The measure of workers is normalized to 1. The skill distribution is taken as exogenous and we denote the fraction of the labor force with skill level  $s_i$  as  $p_i$  with  $\sum_{i=1}^S p_i = 1$ . Jobs are described by their minimum skill requirement,  $y$ . The technology is such that the output produced by a job of type  $y$  with a worker of skill  $s$  is

$$x(s, y) = \begin{cases} y & \text{if } s \geq y \\ 0 & \text{if } s < y \end{cases}.$$

A job is either vacant or filled. When a job of type  $y$  is filled by a worker of skill  $s$ , a wage of  $w(s, y)$  is paid and a cost of  $c_y$  is incurred. That is, the flow value to the firm of filling a job of type  $y$  with a worker of skill  $s$  is  $y - w(s, y) - c_y$ , conditional, of course, on  $s \geq y$ . The corresponding flow value to the worker holding the job is the wage. When a job is vacant, the fixed cost must still be paid so the flow value of a vacancy of type  $y$  is  $-c_y$ . The corresponding flow value to an unemployed worker is  $b$ , which can be interpreted as unemployment compensation and/or the value of not working.

Matches break up (filled jobs become vacant) at the rate  $\delta_y$ , i.e., we assume that job stability varies by job type (this, as well as the dependence of  $c$  on  $y$ , generalizes Albrecht and Vroman, 2002).<sup>6</sup> The flow in the opposite direction is governed by a matching function. Specifically, unemployed workers and vacancies match randomly according to a constant returns to scale matching function

$$m(u, v) = m\left(1, \frac{v}{u}\right)u = m(\theta)u, \text{ where } \theta = \frac{v}{u}$$

with  $m'(\theta) > 0$  and  $d\left(\frac{m(\theta)}{\theta}\right)/d\theta < 0$ .

We use the following notation:

$U(s)$  is the value of unemployment for a worker of skill  $s$

$N(s, y)$  is the value of employment for a worker of skill  $s$  on a job of type  $y$

$V(y)$  is the value of a vacancy of type  $y$

$J(s, y)$  is the value to an employer of filling a job of type  $y$  with a worker of skill  $s$ .

A match will be formed if and only if

$$N(s, y) + J(s, y) \geq U(s) + V(y),$$

and when a match is formed, the wage,  $w(s, y)$ , is given by the Nash bargaining condition,

$$N(s, y) - U(s) = \beta[N(s, y) + J(s, y) - U(s) - V(y)],$$

where  $\beta$  is the exogenously given worker's share of the surplus.

We assume free entry and exit of vacancies, so in equilibrium, there will be at most  $S$  skill requirements:  $y_j = s_j$ ,  $j = 1, \dots, S$ . We define  $\phi_j$  to be the fraction of vacancies requiring skill  $s_j$  and  $\gamma_i$  to be the fraction of the unemployed who have

---

<sup>6</sup>Note that  $\delta$  has a different meaning than in Section 4.

skill  $s_i$ . The unemployment rate,  $u$ , labor market tightness,  $\theta$ , and the fractions  $\{\phi_j\}_{j=1}^S$  and  $\{\gamma_i\}_{i=1}^S$  are the fundamental endogenous variables of the model.

The value functions for filled jobs are

$$\begin{aligned} rN(s_i, s_j) &= w(s_i, s_j) + \delta_j[U(s_i) - N(s_i, s_j)] \\ rJ(s_i, s_j) &= s_j - w(s_i, s_j) - c_j + \delta_j[V(s_j) - J(s_i, s_j)]. \end{aligned}$$

Both of these are conditional on  $s_i \geq s_j$ . The value of unemployment for a worker of skill  $s_i$  is

$$rU(s_i) = b + m(\theta) \sum_{j \leq i} \phi_j \max[N(s_i, s_j) - U(s_i), 0],$$

and the value of a vacancy of type  $s_j$  is

$$rV(s_j) = -c_j + \frac{m(\theta)}{\theta} \sum_{j \leq i} \gamma_i \max[J(s_i, s_j) - V(s_j), 0].$$

Free entry and exit of vacancies implies  $V(s_j) \leq 0$ , with equality if  $\phi_j > 0$ ,  $j = 1, \dots, S$ .

The above expressions imply that a match will be formed if and only if

$$s_j - c_j \geq rU(s_i).$$

and the wage of a worker of skill  $s_i$  on a job requiring skill  $s_j$  is

$$w(s_i, s_j) = \beta(s_j - c_j) + (1 - \beta)rU(s_i).$$

Both of these are conditional on  $s_i \geq s_j$ .

We look for steady-state equilibria. A steady-state equilibrium is a collection of variables  $u$ ,  $\theta$ ,  $\{\phi_j\}_{j=1}^S$ , and  $\{\gamma_i\}_{i=1}^S$  such that (i) the appropriate steady-state conditions hold, (ii) there is free entry and exit of vacancies, i.e.,  $V(s_j) \leq 0$  ( $= 0$  if  $\phi_j > 0$ ), and (iii) matches form iff  $s_j - c_j \geq rU(s_i)$ . Several equilibrium types are possible. For example, one might consider an equilibrium in which workers at each skill level match only with vacancies requiring precisely that skill, i.e., an equilibrium with perfect assortative matching. We refer to this case as equilibrium with *ex post* segmentation. At the other extreme, an equilibrium might entail all possible matches; i.e., a worker of skill  $s_i$  could match with any job of type  $s_j \leq s_i$ . We refer to this case as equilibrium with full cross-skill matching. Intermediate cases, in which some but not all possible matches are formed, are also possible.

The nature of equilibrium depends on the exogenous parameters of the model. If an equilibrium of a particular type exists, e.g., an equilibrium with *ex post* segmentation, then that equilibrium is unique within that class. There may, however,

be multiple equilibria in the sense that equilibria of more than one type can exist simultaneously.

We model KL as a change in the proportions of the labor force in the various skill categories, moving low-skill workers to the medium skill level. Before KL, the labor force has a particular skill distribution and we assume that the market is in a corresponding steady-state equilibrium. We use pre-KL data to calibrate this equilibrium. After KL, the labor force has another skill distribution, with more medium-skill workers and fewer low-skill workers. We simulate the new labor market equilibrium on the basis of the new skill distribution as well as the structural parameters obtained in the calibration. Comparison of the equilibrium outcomes is informative about the equilibrium effects of KL for various worker types. We are particularly interested in the changes in outcomes for individuals who were previously low skilled and currently medium skilled.

A theoretical analysis of the transition path between the equilibria is infeasible. Also, a comparative statics exercise that compares two equilibria can not be translated into a sequence of actions and reactions by individual agents. However, from the above model we can get some idea about the underlying mechanisms. With more medium-skilled workers, the rate at which employers contact them increases. Similarly, the rate at which low-skill workers are contacted decreases. This provides an incentive for employers to create medium-skill jobs rather than low-skill jobs. Labor demand thus adjusts to labor supply. The extent to which this occurs depends on the parameters of the model and on the assumed specification of the production function. Simultaneously, labor market tightness, the transition rates from unemployment to employment by worker-skill level, the unemployment rate, and the wage rates by worker/job-skill combination change. Again the results depend on the model parameters and functions.

It is important to point out that we do not assume that KL has a direct effect on the individual contact rate for a given skill level in a given equilibrium. So in this sense there is no causal “job search assistance” effect on the transition rates to work. However, individuals who are treated in KL qualify for a different set of jobs in the new equilibrium. In addition, since KL affects the proportions of low-skill workers and low-skill jobs, there is an indirect effect on the transition rates to work even for individuals who do not change skill level.

## 5.2 Application to the Knowledge Lift

### 5.2.1 Data for the calibration

The data we use are for 1996, the year before the start of the KL, and are taken from publications from Statistics Sweden. In particular, we use data from AKU, which is the Swedish labor force survey, and from the Lönestatistisk Årsbok, which gives wage data aggregated from firm records (see Statistics Sweden, 1997a, 1997b).

For our calibration, we use the crudest possible set of skill levels, namely, low-skill ( $s_1$ ) and medium-skill ( $s_2$ ). We assume exogenous skill fractions, identifying skill with educational attainment. We take those with less education than a 3-year gymnasium degree, i.e., SUN Codes 1, 2 and 3, to be low-skill and those with a 3-year gymnasium degree (SUN Code 4) and those with less than 3 years of post-gymnasium education (SUN Code 5) to be medium-skill<sup>7</sup>. AKU Table 48 provides observations of the fractions of labor force participants and unemployed participants aged 25-64 by SUN Codes,

SUN Codes	1	2	3	4	5	Total
LF in 100's	4320	4544	12346	5870	5649	32729
U in 100's	398	431	996	448	249	2522
U rates	.092	.095	.081	.076	.044	.077

These imply that  $p_1 = 0.648$  and  $p_2 = 0.352$ . In addition, the fraction of unemployment accounted for by low-skill workers  $\gamma_1$  equals  $\frac{398+431+996}{2522} = 0.724$  (implying that  $\gamma_2 = 0.276$ ), and the average skill-specific unemployment rates are  $u_1 = 0.086$  and  $u_2 = 0.060$ .

We also quantify the exit rates out of unemployment for the two skill groups. For this we use AKU Table 49, giving the elapsed unemployment duration distribution by skill. Our model assumes exponential duration distributions. The exponentiality assumption helps us in two ways. We have data on elapsed, as opposed to completed, durations. The exponential assumption implies that these two distributions, i.e., of elapsed and completed durations, are the same. Second, if  $\xi$  is the median of an  $\exp\{\lambda\}$  distribution, then  $\lambda = \frac{\ln 2}{\xi}$ , i.e., we can use the

---

<sup>7</sup>SUN stands for Swedish Education Level. There are 7 SUN codes. Category 6 is 3 or more years of post-gymnasium education and category 7 is doctoral education. We assume there is no interaction between the labor markets for workers in these higher skill levels and those that we are calling medium skill. We attempted to calibrate a three-skill level model extension, but it appeared that the data are inconsistent with such a model.

median of the elapsed duration distribution to estimate the exponential parameter. As a result, the exit rates out of unemployment for low and medium skills equal 1.899 and 2.163, respectively, in per-year terms.

With an eye on the “Totalt” column of Table 4 of Lönestatistisk Årsbok, we quantify the wages as  $w_{11} = 177,600$ ,  $w_{21} = 186,000$ ,  $w_{22} = 210,000$ , where  $w_{ij}$  is a shorthand for  $w(s_i, s_j)$ .

### 5.2.2 Relation to the microeconomic analyses

It is useful to address some major similarities and differences between the data used in this section and those used in the previous section.

First, the micro data do not provide wage rates but annual income. Moreover, the micro data only allow for an imperfect distinction between full-time and part-time jobs and for a distinction between consecutive spells of employment and non-participation. This makes it difficult to use the micro data to calibrate the equilibrium model, which is why we resort to the wage statistics for the latter. The other aggregate data used in the calibration are taken to be as consistent as possible with each other and with the wage data.

Second, equilibrium models with frictions cannot incorporate the amount of heterogeneity observed in the micro data without becoming intractable. This applies to *ex ante* heterogeneity in individuals as well as to *ex post* heterogeneity in individual outcomes. This is why the present section only distinguishes between a small number of skill types and why the equilibrium only has a small number of possible wage outcomes. The equilibrium model does not incorporate unobserved (to us) heterogeneity and accordingly does not capture selective enrollment into KL. Accordingly, we feel that in the present section, it does not make sense to perform separate analyses for different age or gender categories. The heterogeneity in average outcomes between categories is of a much smaller order than the ignored heterogeneity in outcomes within categories.

We should also recall that the microeconomic analysis includes individuals who are treated in KL but do not raise their skill level from low to medium as a result, for example because the courses they take do not provide marketable skills or the courses lead to an improvement of skills that is insufficient to reach the medium skill level. This means that our microeconomic evaluation may underestimate the average effect of an upgrading to the medium skill level.

Next, we use data from 1996 to calibrate the pre-KL equilibrium, whereas in the microeconomic analysis we use 1994 as the pre-program year. The latter is warranted because of the pre-program dip among those who are treated in KL. However, the aggregate data from 1996 that are used in the calibration

consist mostly of individuals who either are never treated in KL or are not treated immediately after 1996. Using aggregate data from earlier pre-KL years cannot be expected to improve the calibration.

Finally, since we are interested in equilibrium outcomes, and since the labor market in 2000 was clearly not in the post-KL equilibrium, we cannot use data from our latest year in the micro datasets to calibrate the new equilibrium.

### 5.2.3 The calibration of the pre-KL equilibrium

Since we are considering a model with 2 skill levels, there are 2 possible equilibrium configurations, namely

- (i). Cross-skill matching: In this equilibrium, medium-skill workers match with both medium-skill and low-skill vacancies. Low-skill workers match only with low-skill vacancies.
- (ii). *Ex post* segmentation: In this equilibrium, medium-skill workers match only with medium-skill vacancies, and low-skill workers match only with low-skill vacancies.

The large amount of variation in the wage data for medium-skilled workers fits better with the first configuration. Indeed, in Appendix 2, we demonstrate that calibration of the second configuration provides nonsensical results. We therefore base our analysis on the cross-skill matching equilibrium.

The first step in the calibration is to use the exit rates from unemployment and the steady-state conditions to identify the flow parameters, namely,  $\delta_1, \delta_2, \phi_1, \phi_2$ , and  $m(\theta)$ . For low-skill workers, the exit rate from unemployment equals  $m(\theta)\phi_1$ , while for medium-skill workers it equals  $m(\theta)(\phi_1 + \phi_2) \equiv m(\theta)$ . From the values of these exit rates we can therefore uncover  $m(\theta)$ ,  $\phi_1$ , and  $\phi_2$ . We check at this point that  $\phi_i \in [0, 1]$ .

The first steady-state condition is that the flow of low-skill workers into low-skill employment equals the flow of low-skill workers back into unemployment. This can be expressed as

$$\phi_1 m(\theta) \gamma_1 u = \delta_1 e_{11},$$

where  $e_{11}$  is the fraction of the labor force accounted for by employment of low-skill workers in low-skill jobs. Given our estimated exit rates and data on unemployment by skill level, we know the value of the left-hand side of this equation. Further, since  $e_{11} = p_1 - \gamma_1 u$ , we can compute the remaining unknown in this first

steady-state equation, namely,  $\delta_1$ . The second steady-state condition is that the flow of medium-skill workers into low-skill employment equals the corresponding flow from low-skill employment back into unemployment,

$$\phi_1 m(\theta) \gamma_2 u = \delta_1 e_{21}.$$

This condition gives us  $e_{21}$ , the fraction of the labor force accounted for by medium-skill workers employed in low-skill jobs. Next, the flow of medium-skill workers into medium-skill employment equals the corresponding flow from medium-skill employment back into unemployment. That is,

$$\phi_2 m(\theta) \gamma_2 u = \delta_2 e_{22}.$$

We know  $e_{21} + e_{22}$ , i.e., total employment of medium-skill workers. We know  $e_{21}$  from the second steady-state condition, so we have  $e_{22}$ . The third steady-state condition then gives us  $\delta_2$ .

In the second step, we set values for  $b$  and  $r$ . Given the three wage equations

$$w(s_i, s_j) = \beta(s_j - c_j) + (1 - \beta)rU(s_i) \text{ for } s_i \geq s_j$$

and the expression for  $rU(s_1)$ , we can solve for  $s_1 - c_1$ ,  $s_2 - c_2$ ,  $rU(s_1)$ ,  $rU(s_2)$ , and  $\beta$ .

At this point, we need to check that the relevant conditions on these values for a cross-skill matching equilibrium hold, namely,

$$\begin{aligned} s_1 - c_1 &\geq rU(s_2) \\ s_1 - c_1 &\geq rU(s_1) \\ s_2 - c_2 &\geq rU(s_2). \end{aligned}$$

If these are not satisfied then the parameters of the model are inconsistent with this type of equilibrium.

The third step of our calibration strategy is to use the zero-value conditions to recover the cost parameters and the parameters of the matching function. At this point, we need to fix two more parameters. We assume a Cobb-Douglas matching function, so  $m(\theta) = A\theta^\alpha$ , and we choose plausible values for  $A$  and  $\alpha$ . We choose  $\alpha = 0.5$ , with reference to estimates from the empirical literature on matching functions (e.g., Petrongolo and Pissarides 2001). The choice of  $A$  is more arbitrary, but since we have already recovered  $m(\theta)$  from the first step of our procedure, a choice of  $A$  is equivalent to choosing  $\theta$ . Since the numerator of  $\theta$  (i.e., the measure of vacancies) is difficult to quantify, this can be viewed as a normalization. We take  $A = 5$ .



Finally, the zero-value conditions give us  $c_1$  and  $c_2$ . We would naturally like  $c_1 < c_2$  and  $s_1 < s_2$ . We assume that  $b = 80000$  and  $r = 0.05$ .

We now give the results. The equations for the exit rates from unemployment imply

$m(\theta) = 2.163$	$\phi_1 = 0.878$	$\phi_2 = 0.122$
---------------------	------------------	------------------

Putting these values into the steady-state conditions allows us to recover the job-specific exit rates and the skill composition of employment as described above. These are

$\delta_1$	$\delta_2$	$e_{11}$	$e_{21}$	$e_{22}$
.179	.053	.592	.225	.106

Next, from step 2,

$s_1 - c_1$	$s_2 - c_2$	$rU(s_1)$	$rU(s_2)$	$\beta$
190670	244530	167100	182250	.44557

Note that the inequalities required for cross-skill matching are satisfied.

Finally, we solve for  $\theta$ ,  $c_1$ ,  $c_2$ ,  $s_1$ , and  $s_2$ . Given  $m(\theta) = 2.163$ , our choice of  $\alpha$  and  $A$  implies  $\theta = 0.187$  and  $\frac{m(\theta)}{\theta} = 11.56$ . We recover the cost parameters from the zero-value conditions. This gives

$\theta$	$c_1$	$c_2$	$s_1$	$s_2$
.187	542650	1069400	733320	1313930

This solution ranks the cost and productivity parameters in the desired order.

The results can be used to assess the effects of upgrading the skill level of a single low-skill individual to the medium skill level. In a market with a continuum of workers the upgrading has no measurable effects on the other agents, and we restrict attention to the effects on outcomes for the individual under consideration. We may call these the “partial” effects of a skill-raising program. They do not have an empirical counterpart, but as we shall see, they are helpful to understand the equilibrium effects.

For such an individual, the average wage changes from  $w_{11}$  to the mean wage among medium-skilled workers ( $\sum_j e_{2j}w_{2j} / \sum_j e_{2j}$ ), which is an increase of 16086. The average unemployment rate changes from  $u_1$  to  $u_2$ , i.e. it decreases by 2.6 percentage points. This difference is partly due to the fact that medium-skill jobs have a much lower job separation rate than low-skill jobs, and partly due to the

higher exit rate out of unemployment for medium-skill workers. Subsequently, it can be deduced that the average annual income (which is a weighted average of  $b$  and the average wage) increases by 17640. This exceeds the effect on the average wage, despite the fact that income in unemployment is independent of skill level. This is of course due to the dominating reduction of the average amount of time spent unemployed.

#### 5.2.4 Simulation of the KL

Suppose the KL reduces the proportion of low-skill workers from  $p_1 = .648$  to  $.60$  and then to  $.55$  and increases the proportion of medium-skill workers from  $p_2 = .352$  to  $.40$  and then to  $.45$ . We feel that these capture the magnitude of the change in skill levels that the KL could potentially produce. Recall that the other structural parameters are fixed to the values obtained above, i.e.,  $r = 0.05, b = 80000, m(\theta) = 5\theta^5, \delta_1 = 0.179, \delta_2 = 0.053, c_1 = 542650, c_2 = 1069400, s_1 = 733320, s_2 = 1313930$ . In Appendix 3 we derive the equations that must be solved for the simulation.

##### Simulation with $p_1 = .6$ and $p_2 = .4$

$\theta$	$m(\theta)$	$u$	$u_1$	$u_2$	$\gamma_1$	$\gamma_2$	$\phi_1$	$\phi_2$
.18018	2.1224	.078	.092	.056	.71017	.28983	.82873	.17127
$e_{11}$	$e_{21}$	$e_{22}$	$w_{11}$	$w_{21}$	$w_{22}$			
.54458	.22225	.15513	176790	188270	212270			

##### Simulation with $p_1 = .55$ and $p_2 = .45$

$\theta$	$m(\theta)$	$u$	$u_1$	$u_2$	$\gamma_1$	$\gamma_2$	$\phi_1$	$\phi_2$
.17356	2.083	.079	.099	.053	.69537	.30463	.77804	.22196
$e_{11}$	$e_{21}$	$e_{22}$	$w_{11}$	$w_{21}$	$w_{22}$			
.49530	.21698	.20906	175890	190400	214400			

We first discuss the equilibrium effects on aggregate outcomes and then the equilibrium effects at the individual level.

The results suggest that a program like KL that moves workers from low to medium skills can have important equilibrium effects. The first-order equilibrium effect comes via the equilibrium change in job composition. As the fraction of medium-skill workers in the labor force increases, the fraction of vacancies tailored towards those workers increases commensurately. In our simulations, increases in  $p_2$  translate almost one-for-one into increases in  $\phi_2$ . The measure of low-skill workers employed in low-skill jobs ( $e_{11}$ ) also falls by about the same amount.

There is a slight decrease in the measure of medium-skill workers employed in low-skill jobs ( $e_{21}$ ). On the one hand, there are more medium-skill workers; on the other hand, there are fewer low-skill jobs. In our simulations, the second effect dominates slightly. Finally, the increase in the measure of medium-skill workers employed in medium-skill jobs increases by about the same amount as the fraction  $p_2$  of medium-skilled workers does.

The effect of a change in skill composition on *aggregate* unemployment is small. This is partly because the change in labor market tightness is small. The variable  $\theta$  falls slightly, meaning that workers in general take a bit longer to locate a vacancy. However, there are important distributional effects on unemployment across the two skill categories. There are more medium-skill workers, and these workers on average find jobs more quickly and on average retain them longer than low-skill workers do. Unemployment even decreases among the fraction of medium-skill workers who have always been medium-skill, because, even though  $m(\theta)$  falls slightly, there are relatively more medium-skill jobs which are on average kept longer. At the same time, the remaining low-skill workers have more difficulty finding a job than they did before the policy change. The reason is again the shift in job composition – relatively fewer low-skill vacancies are being opened ( $\phi_1$  falls). The fraction of unemployment accounted for by medium-skill workers,  $\gamma_2$ , increases simply because there are now more medium-skill workers.

There is also a clear effect on the distribution of wages. The wages of low-skill workers fall whereas the wages of medium-skill workers increase on both low- and medium-skill jobs. This reflects the change in unemployment values for the two worker types. The value of unemployment among low-skill workers falls because these workers now take longer on average to find a job; the value of unemployment among medium-skill workers increases because these workers now face a better mix of job opportunities.

The top panel of Table 5 summarizes the average wage, employment, and income outcomes, before and after KL, and by skill level. These are subsequently used to quantify the changes in outcomes for those who stay low skilled, those who stay medium skilled, and those whose skills are upgraded, in the bottom panel of the table. For completeness we also list the “partial” effects discussed in Subsection 5.2.3. The last row of Table 5 gives the the effects on the outcomes for the low-skill individuals whose skills are actually upgraded, by comparing them to the post-program outcomes of the individuals whose skills are not upgraded. These are the counterparts of the difference-in-differences and the conditional probit analysis in Section 4. They may be called the average equilibrium treatment effects on the treated.

Table 5: *Equilibrium average treatment effects if KL changes the relative measures of low and medium skilled from 0.65 to 0.55 and from 0.35 to 0.45, respectively.*

outcome measure:	average wage	employment probability	average income
(a) pre-KL, low skill	178	0.914	169
(b) pre-KL, medium skill	194	0.940	187
(c) post-KL, low skill	176	0.901	166
(d) post-KL, medium skill	202	0.947	196
“partial” treatment effect on treated: $b$ minus $a$	16	0.026	18
equil. change for remaining low skilled: $c - a$	-2	-0.013	-3
equil. change for remaining medium skilled: $d - b$	8	0.007	9
equil. change for treated: $d - a$	24	0.033	27
equil. treatment effect: $(d - a) - (c - a)$	26	0.046	29

Note: monetary variables are annual averages in 1000 SEK (1996 level).

The rows in the top panel of the table display the same ranking for each outcome measure:  $c < a < b < d$ . This confirms for each outcome measure that the treated gain most from the program. Those who have always been medium skilled also benefit, whereas the remaining low skilled suffer. Moreover, the equilibrium effects are always on the order of 1.5 to 2 times larger than the “partial” effects. Again, this is because the program generates an equilibrium response of the skill distribution across vacancies towards the higher skill. In this sense, the program has a multiplier effect at the aggregate level.

To what extent do the above treatment effects agree with the microeconomic results in Section 4? The simulated effect on employment is in agreement with the econometric results for young men. However, the finding that the average “partial” effect on wages is positive and bounded from above by the average equilibrium effect on wages is harder to reconcile with the econometric results. We can relate this to the discussion at the end of Section 4. The post-program-

participation year 2000 used in the microeconomic analysis lies only 3 years after the launch of the KL program and is in the middle of the era during which KL ran. In 2000, not enough time had passed to allow the full effects of the program to come to fruition. Indeed, in 2000, individuals whose skills were upgraded may not even have had enough time to leave their post-program dip. Employment effects may reveal themselves earlier than wage effects if the wage-setting institutions do not allow for swift wage adjustments. Also, employers with medium-skill vacancies may initially discriminate against treated workers, while the latter may initially be preferred over low-skill applicants by employers with low-skill vacancies.<sup>8</sup>

Insofar as the economy is able to adjust, our model predicts that the upgrading of the skills of a large fraction of the low-skill work force eventually (in steady-state equilibrium) leads to an economy with more medium-skill jobs and fewer low-skill jobs and with wage increases for those who make the skill upgrade.

## 6 Conclusions

The econometric analysis provides evidence of a positive employment effect for young men. This is remarkable in the sense that (i) training programs have usually been found to be ineffective in raising the employment probability, and (ii) if a program works, then typically it works for women but not for men. It seems that men who were hurt by the recession in the early 1990s are highly represented among the male participants, and they benefited from KL to get back to work – in accordance to the objective of modern adult education. However, we do not find an average income effect. This is at odds with the objective of the program to raise worker skills and thereby worker productivity and wages. The equilibrium calibration evaluation also predicts a positive effect. We conclude that the sample size on individual incomes among treated may be too small. In addition, the econometric analysis may underestimate the average income effect, because of the gain in work experience in the control groups, and/or because the post-program year is too close to the treatment period.

---

<sup>8</sup>Of course, the simulated equilibrium wage effects depend on the way we have specified our model. An equilibrium model along the lines of Heckman, Lochner, and Taber (1998), for example, would give different results. In their model – a Roy model extended to allow for investment in skills and embedded in a dynamic general equilibrium framework with both physical and human capital – a change in the skill distribution of the type envisioned by proponents of the knowledge lift would lead at first to an increase in wage inequality. However, after the transition was completed, wage inequality would decrease. Note, however, that their model rules out any possibility of unemployment.

For women we find no significant effects at all. It seems that female participants were and are similar to non-participants. Perhaps the participation of older female individuals at adult education courses is not directly inspired by labor-market prospects but rather by the consumption value of the education. Jacobson, LaLonde and Sullivan (2003) find that the returns among older individuals depend very strongly on the type of course followed. It would be useful to obtain data with more individual and course characteristics to study the effect of heterogeneity, and to subsequently design a profiling mechanism to sort out those older individuals and courses for whom the effect is likely to be positive. More generally, future research should be based on larger samples of treated, along with better micro data on wages. The present study has also shown that the evaluation of adult education program participation would benefit from data covering a time span of decades rather than years. In particular, with data from more recent post-program years, a more comprehensive picture of the program effects should be possible.

The theoretical analysis of an equilibrium search model with heterogeneity, along with the calibration of the model and the simulation of the policy change, provide some interesting insights into the equilibrium effects of the program. Most notably, the program generates an equilibrium response of the skill distribution of vacancies towards the higher skill. In the simulation, as the fraction of medium-skill workers in the labor force increases at the expense of the fraction of low-skill workers, the fraction of vacancies tailored towards the medium-skill workers increases commensurately, almost one-for-one. For each outcome measure, the treated gain most from the program. Those who have always been medium skilled also benefit, whereas those who remain low skilled suffer. Moreover, the equilibrium effects are always on the order of 1.5 to 2 times larger than the “partial” effects.

## References

- Albrecht, J.W. and S.B. Vroman (2002), “A matching model with endogenous skill requirements”, *International Economic Review*, 43, 283–305.
- Axelsson, R. and O. Westerlund (1999), “Participants in the Knowledge Lift in autumn 1997 and comparison groups: labor market and income conditions before education”, Working paper, Umeå University, Umeå (*in Swedish*).
- Bergemann, A., B. Fitzenberger, and S. Speckesser (2004), “Evaluating the dynamic employment effects of training programs in East Germany using conditional difference-in-differences”, Working paper, University of Mannheim, Mannheim.
- Björklund, A., P.A. Edin, P. Fredriksson, and A. Krueger (2004), “Education, equality, and efficiency – An analysis of Swedish school reforms during the 1990s”, Working paper, IFAU, Uppsala.
- Blundell, R., M. Costa Dias, C. Meghir, and J. Van Reenen (2004), “Evaluating the employment impact of a mandatory job search assistance program”, *Journal of the European Economic Association*, 2, forthcoming.
- Carling, K., B. Holmlund, and A. Vejsiu (2001), “Do benefit cuts boost job findings? Swedish evidence from the 1990s”, *Economic Journal*, 111, 766–790.
- Ekström, E. (2003), “Essays on inequality and education”, Working paper, Uppsala University, Uppsala.
- Fay, R.G. (1996), “Enhancing the effectiveness of active labour market policies: Evidence from programme evaluations in OECD countries”, Working paper, OECD, Paris.
- Friberg, N. (2000), “Thematic review on adult learning: Sweden”, Working paper, OECD, Paris.
- Heckman, J.J., H. Ichimura, J.A. Smith, and P. Todd (1998), “Characterizing selection bias using experimental data”, *Econometrica*, 65, 1017–1098.
- Heckman, J.J., R.J. LaLonde, and J.A. Smith (1999), “The economics and econometrics of active labor market programs”, in O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics, Volume III*, North-Holland, Amsterdam.
- Heckman, J.J., L. Lochner, and C. Taber (1998), “General equilibrium treatment effects: a study of tuition policy”, *American Economic Review*, 88, 381–386.
- Heckman, J.J. and J.A. Smith (1999), “The pre-program earnings dip and the

- determinants of participation in a social program: implications for simple program evaluation strategies”, *Economic Journal*, 108, 313–348.
- Jacobson, L., R.J. LaLonde, and D. Sullivan (2003), “Should we teach old dogs new tricks? The impact of community college retraining on older displaced workers”, Working paper, Federal Reserve Bank, Chicago.
- Larsson, L. (2003), “Evaluation of Swedish youth labour market programmes”, *Journal of Human Resources*, 38, 891–927.
- Lise, J., S. Seitz, and J. Smith (2002), “Equilibrium policy experiments and the evaluation of social programs”, Working paper, Queen’s University, Kingston.
- Magnac, T. (2004), “Panel binary variables and sufficiency: generalizing conditional logit”, *Econometrica*, 72, 1859–1876.
- Martin, J.P. and D. Grubb (2001), “What works and for whom: a review of OECD countries’ experiences with active labour market policies”, *Swedish Economic Policy Review*, 8, 9–56.
- Ministry of Education (1998), “Adult education and life long learning. the situation before and during the first year of the Knowledge Lift”, Working paper, Ministry of Education, Sweden, Stockholm (*in Swedish*).
- Ministry of Industry (2001), “Sweden’s action plan for employment”, Working paper, Ministry of Industry, Sweden, Stockholm.
- National Agency for Education (1999), *The Adult Education Initiative*, National Agency for Education, Stockholm.
- OECD (2004), *Employment Outlook 2004*, OECD, Paris.
- Petrongolo, B. and C.A. Pissarides (2001), “Looking into the black box: a survey of the matching function”, *Journal of Economic Literature*, 39, 390–431.
- Richardson, K. and G.J. van den Berg (2001), “The effect of vocational employment training on the individual transition rate from unemployment to work”, *Swedish Economic Policy Review*, 8, 175–213.
- Skolverket (2001), “The Knowledge Lift and upper secondary adult education Fall 2000”, Working paper, Skolverket, Stockholm (*in Swedish*).
- Statistics Sweden (1997), “Labor force survey, yearly average 1996”, Working paper, Statistics Sweden, Örebro (*in Swedish*).
- Statistics Sweden (1997), *Wage statistics yearbook 1996*, Statistics Sweden Publication Service, Örebro (*in Swedish*).



Statskontoret (1999), “How much does the Knowledge Lift cost?”, Working paper, Statskontoret, Stockholm (*in Swedish*).

Stenberg, A. (2003), “An evaluation of the adult education initiative relative labor market training”, Working paper, Umeå University, Umeå.

# Appendix

## Appendix 1. Miscellaneous estimation results

Table 6: *Estimates of the average treatment effects on the treated, concerning participation in KOMVUX in 97-II, using 1994 and 2000 as pre- and post-program years and annual income as outcome measure (in 1000 SEK in the year 2000; trimmed income data).*

type of individual	ATET	(standard error)	$n_{treated}$	$n_{control}$
Male, aged 25–40	-8.0	(13)	25	3729
Female, aged 25–40	6.3	(8)	46	4427
Female, aged 41–55	4.1	(9)	33	7162

## Appendix 2. Calibration of the *ex-post* segmentation equilibrium

In an *ex post* segmentation equilibrium, there is no steady-state condition for the flows of medium-skill workers in and out of low-skill jobs. We therefore assume that  $\beta = 0.5$ . The inequalities on  $U$  are now:

$$\begin{aligned} s_1 - c_1 &< rU(s_2) \\ s_1 - c_1 &\geq rU(s_1) \\ s_2 - c_2 &\geq rU(s_2). \end{aligned}$$

The exit rates for this case are

$m(\theta)\phi_1 = 1.899$
$m(\theta)\phi_2 = 2.163$

These imply:

$m(\theta) = 4.062$	$\phi_1 = .468$	$\phi_2 = .532$	
$\delta_1$	$\delta_2$	$e_{11}$	$e_{22}$
.179	.139	.592	.331

Step 2 yields:

$s_1 - c_1$	$s_2 - c_2$	$rU(s_1)$	$rU(s_2)$
188100	207480	167100	188520

Note that  $s_1 - c < rU(s_2)$ , which is consistent with ex post segmentation.

Finally, in this case,  $m(\theta) = 4.062$ , so  $\theta = .660$  and  $\frac{m(\theta)}{\theta} = 6.15$ . The zero-value conditions imply:

$\theta$	$c_1$	$c_2$	$s_1$	$s_2$
.660	204310	85203	392410	292683

Clearly, the cost and productivity parameters are in the wrong order.

### Appendix 3. Simulation equations

The two steady-state equations each equate the flow into and out of unemployment for one of the skill levels. These two steady-state equations can be expressed as

$$\phi_1 m(\theta)(1 - \gamma_2)u = \delta_1(p_1 - (1 - \gamma_2)u)$$

$$m(\theta)\gamma_2 u \left( \frac{\phi_1}{\delta_1} + \frac{(1 - \phi_1)}{\delta_2} \right) = p_2 - \gamma_2 u$$

These two equations give  $u$  and  $\gamma_2$  in terms of  $\theta$  and  $\phi_1$  (and, of course,  $\phi_1 + \phi_2 = \gamma_1 + \gamma_2 = 1$ .) The second of these equations is derived by combining the two steady-state equations for medium-skill workers, i.e., for flows into and out of low- and high-skill jobs.

In cross-skill matching equilibrium, the unemployment values are

$$rU(s_1) = \frac{bR_1 + m(\theta)\phi_1\beta S_1}{R_1 + m(\theta)\phi_1\beta}$$

$$rU(s_2) = \frac{bR_1R_2 + \beta m(\theta)[\phi_1R_2S_1 + \phi_2R_1S_2]}{R_1R_2 + \beta m(\theta)(\phi_1R_2 + \phi_2R_1)}$$

where  $S_1 = s_1 - c_1$ , and  $S_2 = s_2 - c_2$ ,  $R_1 = r + \delta_1$ , and  $R_2 = r + \delta_2$ .

The vacancy value equations for this equilibrium type are

$$rV(s_1) = -c_1 + \frac{m(\theta)}{\theta} \left\{ \gamma_1 \left[ \frac{(1 - \beta)[S_1 - rU(s_1)]}{R_1} - V(s_1) \right] + \gamma_2 \left[ \frac{(1 - \beta)[S_1 - rU(s_2)]}{R_1} - V(s_1) \right] \right\}$$

$$rV(s_2) = -c_2 + \frac{m(\theta)\gamma_2}{\theta} \left[ \frac{(1 - \beta)S_2 - rU(s_2)}{R_2} - V(s_2) \right]$$

Setting  $V(s_1) = V(s_2) = 0$  and substituting for the unemployment values gives

$$\frac{c_1 R_1 \theta}{(1 - \beta)m(\theta)} = \frac{\gamma_1(S_1 - b)R_1}{R_1 + m(\theta)\phi_1\beta} + \gamma_2\left(\frac{(S_1 - b)R_1R_2 + \beta m(\theta)\phi_2R_1(S_1 - S_2)}{R_1R_2 + \beta m(\theta)(\phi_1R_2 + \phi_2R_1)}\right)$$

$$\frac{c_2 R_2 \theta}{(1 - \beta)m(\theta)} = \gamma_2\left[\frac{(S_2 - b)R_1R_2 + \beta m(\theta)\phi_1R_2(S_2 - S_1)}{R_1R_2 + \beta m(\theta)(\phi_1R_2 + \phi_2R_1)}\right]$$

These two equations along with the two steady-state equations are the equations that must be solved for the equilibrium. After solving for the equilibrium, the wages can be found by using the wage equations

$$w(s_1, s_1) = \beta S_1 + (1 - \beta)rU(s_1)$$

$$w(s_2, s_1) = \beta S_1 + (1 - \beta)rU(s_2)$$

$$w(s_2, s_2) = \beta S_2 + (1 - \beta)rU(s_2)$$