

IZA DP No. 1095

**Regional Treatment Intensity
as an Instrument for the Evaluation
of Labour Market Policies**

Markus Frölich
Michael Lechner

March 2004

Regional Treatment Intensity as an Instrument for the Evaluation of Labour Market Policies

Markus Frölich

*SIAW, University of St. Gallen
and IZA Bonn*

Michael Lechner

*SIAW, University of St. Gallen,
CEPR, ZEW and IZA Bonn*

Discussion Paper No. 1095
March 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Regional Treatment Intensity as an Instrument for the Evaluation of Labour Market Policies*

The effects of active labour market policies (ALMP) on individual employment chances and earnings are evaluated by nonparametric instrumental variables based on Swiss administrative data with detailed regional information. Using an exogenous variation in the participation probabilities across fairly autonomous regional units (cantons) generated by the federal government, we identify the effects of ALMP by comparing individuals living in the same local labour market but in different cantons. Taking account of small sample problems occurring in IV estimation, our results suggest that ALMP increases individual employment probabilities by about 15% in the short term for a weighted subpopulation of compliers.

JEL Classification: C14, J68

Keywords: local average treatment effect, active labour market policy, state borders, geographic variation, weak instruments, small sample problems of IV, Switzerland, Fuller estimator

Corresponding author:

Markus Frölich
Swiss Institute for International Economics and Applied Economic Research (SIAW)
University of St. Gallen
Bodanstr. 8
9000 St. Gallen
Switzerland
Email: Markus.Froelich@unisg.ch

* Financial support from the Swiss National Science Foundation (projects 4043-058311 and 4045-050673) and the Grundlagenforschungsfonds HSG (project G02110112) is gratefully acknowledged. Part of the data originated from a database generated for the evaluation of the Swiss active labour market policies together with Michael Gerfin. We are grateful to the Swiss State Secretariat for Economic Affairs (*seco*; *Arbeitsmarktstatistik*) and the Swiss Federal Office for Social Security (*BSV*) for providing the data and to Dragana Djurdjevic, Michael Gerfin and Heidi Steiger for their substantial input in preparing them. We thank Blaise Melly, Ruth Miquel, Jeff Smith, Heidi Steiger and seminar participants in Essen, London (EC2-conference at cemmap), Madrid (CEMFI) and Vienna for helpful comments, in particular Manuel Arellano, Martin Browning, Kevin Denny, Horst Entorf, Jean-Marc Robin, Coen Teulings and Ernesto Villanueva. The usual disclaimer applies.

1 Introduction

In many empirical applications, geographic borders are exploited as an instrumental variable to uncover the effects of policy interventions. Card and Krueger (1994), Holmes (1998), Black (1999) and Pence (2003) use the U.S. state or district borders to estimate the effects of the minimum wage on employment, of right-to-work laws on manufacturing activity, of school quality on housing prices and of foreclosure laws on mortgage loan size, respectively. Policies change abruptly as one crosses a state border, but the economic environment changes only little within areas close to it. In other words, crossing the border changes the impact of the policy or the likelihood of being subject to the policy, but has no direct effect on individual outcomes in the absence of the policy. Thus, the border acts like an instrumental variable. If there are many state borders, separate effects are identified for each border, because – without further homogeneity assumptions - the instrument is valid only locally. The standard approach consists in specifying a linear IV model to implicitly aggregate these different heterogeneous local effects into a single parameter. In this paper, we introduce a different, *nonparametric* approach for aggregating local effects and for coping with weak instruments. The latter problem is of particular concern in many applications where the policy change accompanied with crossing a state border affects only a fraction of the population. The proposed estimator is used to evaluate Swiss active labour market policies.

Switzerland is a small country with fairly autonomous administrative regions and runs an extensive active labour market policy to counteract unemployment. With short commuting times and a good transport infrastructure, local labour markets are integrated across regional borders and individuals residing on opposite sides close to the border essentially live in the same economic environment. Opportunities for wage arbitrage through relocating instead of commuting hardly exist. When an employed person becomes unemployed, however, a regional difference in active labour market programmes matter. Although the labour market programmes are largely similar throughout the country, treatment incidence is not. This pattern can be explained by regional minimum quota requirements. They were enacted by the Swiss federal government to speed up the local implementation of the reform. As a result the probability of participating in active labour market programmes varies between regions. This exposes individuals within the same local

labour market to different treatment probabilities when they become unemployed. Our identification strategy consists in identifying neighbourhoods on both sides of a regional border that belong to the same local labour market. Living on one or the other side of the border acts as an instrument for estimating average effects of active labour market programmes. The conditions for identification are discussed in detail in the paper. To our knowledge, such an approach has not yet been used to evaluate labour market programmes.

In applications of this kind, when allowing for heterogeneous effects, two problems are likely to occur. First, there might be many regional borders with corresponding treatment effects, which are all very imprecisely estimated because of the small number of observations in the border region. Second, the imprecision of the estimates is often exacerbated by weak instruments and the well-known poor finite sample properties of conventional IV estimators. In our application, for example, the weakness of the instrument is visible in that the treatment probability often changes by less than 5 percentage points across the border. To cope with these problems, the paper advances two methodological contributions: We propose a way to deal (1) with weak instruments in this specific context, and (2) to aggregate the separate local effects into a single effect to discern the main pattern from the many noisy estimates for the different borders.

The first contribution regards the choice of the IV estimator in estimating treatment effects with a potentially weak *binary* instrument, which has not been analysed specifically in the literature so far. Although the Wald (1940) estimator is used in almost all empirical applications when the instrument is binary, it is susceptible to weak instruments and can perform poorly in finite samples. In our Monte Carlo simulations, we find that the estimator proposed by Fuller (1977) is superior to the Wald estimator in this situation.

The second contribution is related to summarize the effects coming from a number of noisy estimates. Examining the effects separately for each regional border may provide useful information on their heterogeneity. More often, however, it will lead to the conclusion that all estimates are insignificant. Drawing any conclusions from the statistical analysis then requires aggregation of the various estimates to obtain a summary measure. We propose a nonparametric and versatile approach to combine the different effects to an *aggregated* local average treatment effect. The Fuller estimator seems to be particularly suited for its estimation, because its finite-sample moments exist.

Applying this approach, we estimate the aggregated local average treatment effect of Swiss active labour market policies on subsequent employment chances. This is the treatment effect for compliers, i.e. for the specific subpopulation that is responsive to changes in the minimum quota. We find positive and significant aggregate effects on individual employment probabilities of about 10 percentage-points.

The paper is organised as follows. In Section 2, the Swiss active labour market policies and the origins of the regional variation in treatment intention intensity are described. Section 3 introduces the aggregated local average treatment effect. Section 4 describes the implementation for Switzerland, and Section 5 gives the estimation results. Section 6 concludes. Several appendices contain additional material.

2 The Swiss labour market and active labour market policies

2.1 Regional employment offices, unemployment insurance, and active labour policies

Until the recession of the early 1990s, unemployment was very low in Switzerland, a small country with 26 different administrative regions, called *cantons*. With the recession, the unemployment rate rose rapidly to 5% and triggered a comprehensive revision of the federal unemployment insurance act. This revision, which became effective partly in January 1996 and partly in January 1997, introduced active labour market programmes (ALMP) on a large scale. Although different in some details, the main components of the Swiss ALMP can be found in various programmes in the USA, UK and Germany as well. Programmes can be grouped into three categories: a) *Training* programmes range from basic skills courses, language courses, computer courses to specific work-related training,¹ with a usual duration of 1 week to 3 months and are carried out by private education providers; b) *Employment programmes* are temporary job creation schemes and consist of provisional or project work for about 3-6 months in the public administration or other public institutions (hospitals, old people's homes, nursing homes, schools, and kindergartens) or in private not-for-profit institutions (e.g. charities, cultural, environmental, recycling organisations). c) *Temporary wage subsidies* (*Zwischenverdienst*) is a programme rather unique to Switzerland and is a subsidy for temporary jobs in the regular labour market. Whereas employment programmes have to be extraordi-

¹ Business and trade training, manufacturing and technical training, etc.

nary, i.e. not competing with regular jobs in the private or public sector, temporary wage subsidies are for jobs in the regular labour market. The subsidy consists of 80% of the pay difference to the previous earnings and is paid to the unemployed person. The subsidy can be granted for up to one year, with an average duration of about 4 months.² During participation in ALMP the unemployed must continue her job search activities and accept any reasonable job offer.

With the reform, benefit entitlement was prolonged to two years, but benefit payments were made conditional on willingness to participate in labour market programmes. This *activation principle* empowered the caseworker to assign an unemployed at any time to any programme provided that participation is expected to be beneficial to her employability. Non-cooperation by the unemployed person can be (and often is) sanctioned through the suspension of benefits.³

Another element of the reform was the consolidation of the 3000 municipal unemployment offices into about 150 regional employment offices (REO), supervised by 26 cantonal centres (CC). These centres contract private and public organisations for providing programmes, compile a catalogue of courses and programmes offered by the contracted providers and seek to ensure that a sufficient number of programme places can be offered as demanded by the REO. The CC have substantial autonomy in administrating and implementing active labour market programmes. The REOs are geographically organised, each REO serving several municipalities. For each unemployed there is one unique REO defined by her place of residence. They cannot change their assigned REO other than by moving to another municipality. Exceptions are the city centres of Zurich and Geneva, which are served by several REOs.

2.2 The minimum quota as an instrument

Since Swiss cantons enjoying a high degree of autonomy might have been slow to implement the new federal policies, they were obliged by law to fill a minimum number of places in active labour market pro-

² Officially, temporary wage subsidies is not labelled an active labour market programme. More details on Swiss active labour market policies can be found in Gerfin and Lechner (2002), Gerfin, Lechner and Steiger (2002), and Lalive, van Ours and Zweimüller (2000).

³ The obligation to be willing to participate in ALMP starts from the first day of unemployment.

grammes per year. For the year 1998, the minimum was 25'000 year-places⁴ and was distributed across the cantons according to the formula

$$12'500 \cdot (\text{population share}_{1996} + \text{unemployment share}_{1996}),$$

where population share is the fraction of the population living in the respective canton as of 1996 and unemployment share is the average number of unemployment benefit recipients in the period April 1996 to March 1997 in the respective canton relative to the total for Switzerland.

The costs of active labour market programmes and of their administration generally are borne by the federal unemployment insurance funds.⁵ The cantons pay a very small lump sum contribution of 3000 Swiss Francs (CHF) per year-place for their assigned minimum quota. They can reduce this lump sum payment by up to 25% if the average unit costs of the purchased programme slots are below the national average within defined programme categories. No financial contribution has to be paid for places filled beyond the required minimum.⁶ On the other hand, cantons which fill less than the required minimum number of year-places, have to compensate the federal unemployment insurance funds with 20% of the unemployment benefits paid to those persons to whom no ALMP could be offered.⁷ Hence, there are financial and political incentives for the cantons to meet their quota. In fact, they were encouraged to provide even more ALMP places.

The formula for the calculation of the quota for 1998 was codified in November 1996, and in October 1997 the minimum quotas for 1998, as given in Table 2.1, were proclaimed.⁸ This formula for the computation of the minimum quota induced regional variation in treatment (participation) intention. Relative to the number of unemployed persons, the quota was rather high in cantons with a low unemployment rate in 1996 because 50% of the quota was distributed according to the population share.

⁴ A year-place corresponds to 220 programme days.

⁵ Art. 92 AVIG (Arbeitslosenversicherungsgesetz), Art. 122a, 122b AVIV (Arbeitslosenversicherungsverordnung).

⁶ Art. 72c (AVIG) and Art. 98b (AVIV).

⁷ Art. 72a (AVIG) and Art. 98b (AVIV).

⁸ Art. 72b (AVIG), Art. 99, 99a (AVIV). Regulation in *Amtliche Sammlung AS 1996 3071, AS 1997 2415*. See also Stauffer, 1998.

Consider the situation of the management of the cantonal and the regional employment offices in the beginning of 1998 for planning their strategy in providing active labour market programmes. In the fourth and fifth column in Table 2.1, the number of registered unemployed in January 1998 and the ratio of the quota to the number of unemployed in January 1998 is given. Suppose the management of the cantonal employment offices forecasted that the number of unemployed would remain constant during the whole year. Then, in cantons such as Uri (UR), Schwyz (SZ), Obwalden (OW), Nidwalden (NW), Appenzell (AR&AI), Glarus (GL) and StGall (SG) with high ratios of the quota to the number of unemployed, the management was under pressure to make sure that many of the relatively few unemployed persons will be assigned to active labour market programmes. In the cantons Zurich (ZH), Solothurn (SO), Schaffhausen (SH), Ticino (TI), Vaud (VD), Neuchâtel (NE), Geneva (GE) and Jura (JU), on the other hand, the relative quota was much lower and the administration was under less pressure to fill this quota. Probably though, the management did not assume that the number of unemployed would remain constant throughout 1998 and their forecasts may have varied between the cantons. Indeed, the number of unemployed decreased by about 30% during the year and this decrease differed between the cantons: from -22% in Zug (ZG) to -66% in AppenzellInner-Rhodes (AI). These differential developments even exacerbated the differences in the quota per unemployed, as cantons where there were relatively few unemployed in January (relative to the quota) experienced larger decreases in the number of unemployed, than cantons where there were relatively many unemployed. Due to these developments, the ratios of the quota to the number of unemployed at the middle or at the end of 1998 (column six in Table 2.1) were even more pronounced. If the cantonal authorities forecasted these trends even roughly, the pressure on those cantons with a high quota per unemployed (with respect to the January figures) was even larger, while it was even less in cantons with a low quota in January. Hence, the *quota per unemployed* in January 1998 indicates the intensity of the pressure to which the local administrations were subjected to assign a sufficient number of unemployed persons to labour market programmes.

Table 2.1: Minimum quotas and number of unemployed

Canton	Minimum quota ^a		Number of unemployed	Quota per unemployed ^b	Quota per unemployed ^c	Realised places ^d	Number of unemployed	Places per unemployed ^e
	1997	1998	Jan 1998	Jan 1998	Dec 1998	total 1998	average 1998	average 1998
Total	25'000	25'000	182'492	13.7	20.1	26'934	139'658	19.3
ZH	4'258	4'325	33'802	12.8	18.0	3'976	27'985	14.2
BE	2'947	2'966	19'591	15.1	25.4	3'665	14'151	25.9
LU	1'000	1'040	6'885	15.1	23.4	1'187	4'967	23.9
UR	64	89	394	22.6	35.0	83	244	33.9
SZ	342	370	1'739	21.3	29.3	533	1'228	43.4
OW	60	75	273	27.5	43.1	56	200	27.8
NW	76	90	381	23.6	46.4	81	263	30.9
GL	111	119	636	18.7	29.2	110	405	27.3
ZG	283	288	1'737	16.6	21.3	305	1'480	20.6
FR	841	805	5'256	15.3	22.1	1'319	4'023	32.8
SO	743	773	6'908	11.2	21.2	820	4'536	18.1
BS	712	685	4'926	13.9	21.2	812	3'855	21.1
BL	774	758	4'740	16.0	27.4	805	3'521	22.9
SH	249	242	2'091	11.6	18.2	285	1'527	18.6
AR	117	142	633	22.4	45.1	118	363	32.5
AI	15	28	112	25.0	73.7	8	56	15.0
SG	1'311	1'370	7'899	17.3	25.1	1'146	6'079	18.8
GR	369	478	3'172	15.1	24.4	433	2'230	19.4
AG	1'629	1'697	10'411	16.3	23.8	1'859	8'276	22.5
TG	656	694	4'742	14.6	23.6	751	3'455	21.7
TI	1'514	1'445	12'383	11.7	16.6	1'828	8'844	20.7
VD	2'833	2'669	21'758	12.3	16.5	2'914	17'885	16.3
VS	1'246	1'194	9'197	13.0	18.8	1'258	5'710	22.0
NE	715	652	5'449	12.0	15.6	1'036	4'513	23.0
GE	1'875	1'750	15'277	11.5	15.1	1'219	12'607	9.7
JU	260	256	2'100	12.2	23.6	327	1'255	26.1

Note: Abbreviations of cantons: ZH: Zurich, BE: Bern, LU: Lucerne, UR: Uri, SZ: Schwyz, OW: Obwalden, NW: Nidwalden, GL: Glarus, ZG: Zug, FR: Fribourg, SO: Solothurn, BS: Basel-Town, BL: Basel-Country, SH: Schaffhausen, AR: AppenzellOuter-Rhodes, AI: AppenzellInner-Rhodes, SG: StGall, GR: Grisons(Graubünden), AG: Argovia(Argau), TG: Thurgovia(Thurgau), TI: Ticino, VD: Vaud, VS: Valais(Wallis), NE: Neuchâtel, GE: Geneva, JU: Jura.

^a) The minimum quota is the minimum number of 'year-places' to be provided by the canton. A year-place corresponds to 220 programme days.

^b) Minimum quota divided by the number of registered unemployed in *January* 1998 (and multiplied by 100).

^c) Minimum quota divided by the number of registered unemployed in *December* 1998 (and multiplied by 100).

^d) Realised places contain only courses, employment programmes and internships (Berufspraktika). Other smaller programmes are not included.

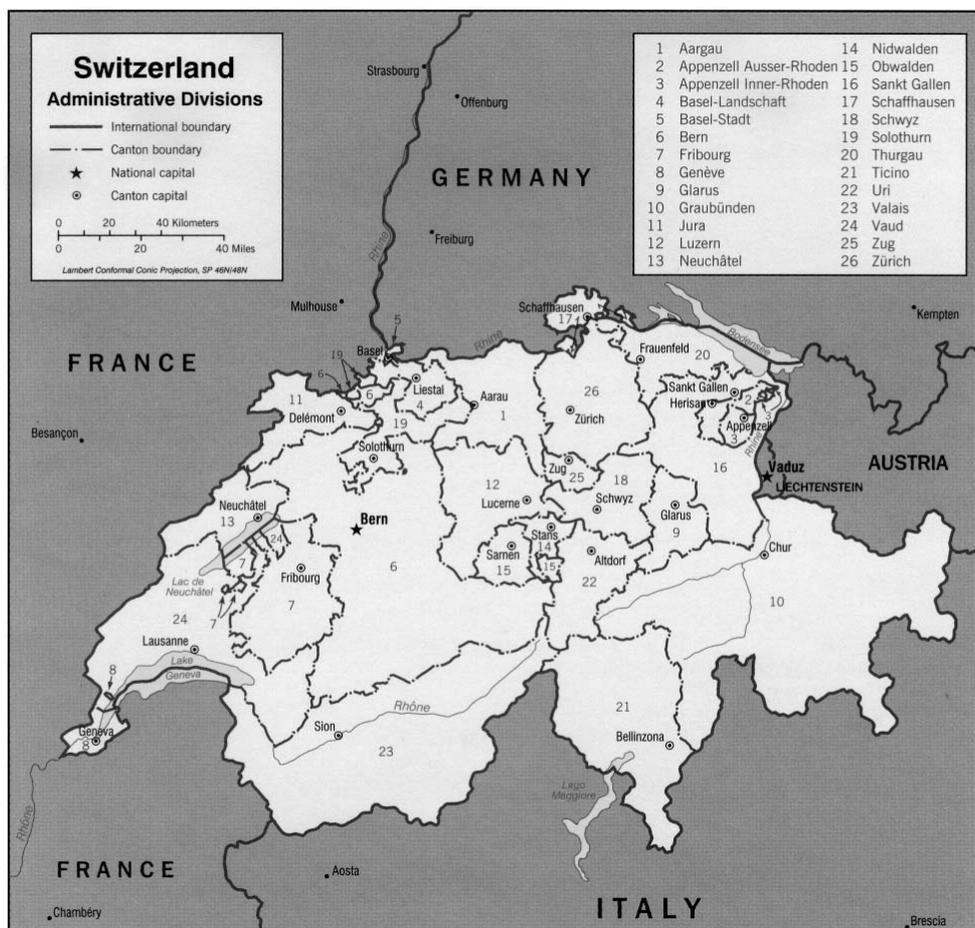
^e) Realised programme places in 1998 divided by the average number of unemployed in 1998 (and multiplied by 100).

Source: Jonathan Gast, seco, Arbeitsmarktstatistik; own calculations.

Columns seven to nine of Table 2.1 show that this measure of treatment intention indeed is correlated with subsequent treatment incidence during the year 1998. Column seven gives the number of year-places that actually were filled in the year 1998, while column eight displays the average number of unemployed in 1998. Column nine shows the actual extent of treatment per unemployed as the ratio of the two previous numbers. The correlation between treatment intention (column 5) and actual treatment incidence (column 9)

is 0.53, thus indicating that the quota indeed induced a higher treatment incidence in cantons with a high relative quota.

Figure 2.1: The cantonal borders of Switzerland



Note: http://www.kyte.de/mygeo/karten/switzerland/landkarte_schweiz.jpg.

While the formula for the calculation of the minimum quota indeed generated a regional variation in treatment intensity, the quota per unemployed is unlikely to be a valid instrument per se, as it is related to the unemployment rate in 1996. The quota per unemployed, however, might be a valid instrument locally if we restrict the analysis to confined areas along cantonal borders. This identification strategy is described in the next section.

3 IV identification with regional variation in treatment intention intensity

Regional variation in treatment intention intensity (i.e. the quota per unemployed) is a candidate instrumental variable for identifying the effects of actual treatment receipt. Since the minimum cantonal quotas are determined by federal law on the basis of the past labour market situation, they are not endogenously chosen according to the preferences of local authorities. The extent to which labour market programmes are finally provided is subject to different regional perceptions in the cantonal administrations about the desirability of ALMP and to the number of available unemployed. Nevertheless, a strong impact of the minimum quota on the share of unemployed persons assigned to labour market programmes has been noted from Table 2.1. However, as argued above, regional variation per se cannot be used for identification as the necessary exclusion restriction is unlikely to be satisfied. In confined regions along the internal administrative boundaries, however, the exclusion restriction should be valid.

The identification is based on the concept of an economically *integrated local labour market* that is divided by an administrative border. A local labour market is considered as integrated if different job opportunities are valued irrespective of the location of residence. In other words, all relevant employment opportunities can be reached within convenient commuting distance (e.g. half an hour) from both sides of the border, such that the choice of workplace location and the choice of residence are not immediately tied. Switzerland, with its numerous administrative borders (Figure 2.1) and a very good commuting infrastructure, is a candidate country for finding such local labour markets. Nevertheless, the methods developed below apply more generally, because in many countries administrative borders often pass through densely populated areas.

Within a local labour market, the different treatment intention intensities on the two sides of the border identify the effects of participation in ALMP on the subsequent labour market career. Two persons residing on opposite sides of the border live in the same economic environment, but when becoming unemployed they have to attend different regional employment offices. This will affect their probability of being assigned to ALMP. The REOs pursue different re-integration strategies, which are partly influenced by the minimum quota the canton has to fulfil. REOs in cantons with an ambitious quota per unemployed will

assign earlier and more persons to programmes than cantons with a lower relative quota. This difference in the treatment probabilities identifies the treatment effect, which is a local average treatment effect (LATE) as defined by Imbens and Angrist (1994). This is further discussed in Section 3.1 for a single local labour market. Section 3.2 examines the situation with multiple local labour markets and introduces the aggregated local average treatment effect. Section 3.3 discusses estimation and the problem of weak instruments. While Section 3 discusses the approach and the assumptions in general, Section 4 describes the particular implementation for Switzerland.

3.1 Local average treatment effect with two administrative regions

Consider a population of unemployed living in a local labour market as defined above and being eligible to take part in active labour market programmes. Let Z denote the instrument *quota per unemployed*, and Z_i be the value of the instrument in the region where person i resides. Consider first a situation with only two administrative regimes, i.e. a single internal border. Hence, $Z_i \in \{z', z''\}$ can take only two different values. Suppose that $z'' \geq z'$. In this setting, the IV approach simplifies to comparing those individuals living to the left of the border with those living to the right of the border, conditional on living close to the border. Individuals who live further apart from the border are outside the local labour market and are not included in the comparison because the IV assumptions are unlikely to be valid for them. Those living in the same local labour market have the same employment opportunities but are subjected to different risks of being assigned to labour market programmes when becoming unemployed. Let D_i indicate whether an unemployed participated in a programme ($D_i=1$) or not ($D_i=0$). Let $D_{i,z}$ denote the potential participation status if the level of the instrument were externally set to z . With the instrument taking only two different values, the potential participation variable $D_{i,z}$ defines four different types of individuals: The always-participants, the never-participants, the compliers and the defiers. Let these different types be denoted by $T_i \in \{a, n, c, d\}$. The always-participants would be assigned to a programme in both cantons. The never-participants would in neither canton be assigned to a programme. The compliers are those who are assigned

to ALMP in the canton with the higher quota per unemployed z'' , but not in the canton with the lower quota z' . For the defiers, this pattern is reversed.

Let Y_i denote an outcome variable of interest (e.g. earnings, employment status) for an individual i and denote her potential outcomes by $Y_{i,z}^d$ for $d \in \{0,1\}$ and $z \in \{z', z''\}$. Define $Y_{i,z} \equiv Y_{i,z}^{D_i(z)}$ as the outcome that would be observed if z were fixed externally. The potential outcomes of interest are $Y_i^d \equiv Y_{i,Z_i}^d$ where d is fixed externally without a change in Z . Since these potential outcomes might be confounded with the participation decision, the causal effect of labour market programmes cannot be inferred directly by simply comparing sample means.

Under conditions discussed below, the treatment effect for the subpopulation of *compliers* is defined as:

$$E[Y^1 - Y^0 | T = c] = \frac{E[Y | Z = z''] - E[Y | Z = z']}{E[D | Z = z''] - E[D | Z = z']}. \quad (1)$$

This local average treatment effect (LATE, Imbens and Angrist, 1994) is the effect of the treatment on those individuals who would switch their treatment status if the value of their instrument would be changed exogenously. Hence, the compliers are those who would not participate in treatment if being subject to the labour market policy in the one canton, but participate if being subject to the labour market policy in the neighbouring canton. As the minimum quota is one of the causes of the different treatment incidence between the cantons, the compliers are the marginal group being induced to enter in treatment due to the differing quotas. Imbens (2001) gives conditions for the identification of a LATE by an instrumental variable. In our setting these general conditions translate to the following specific requirements:⁹

1) Regional treatment intention does not affect the non-treatment outcome: $Y_i^0 \perp\!\!\!\perp Z_i | T_i$

⁹ The following instrumental variables assumptions are exactly identical with the formulae in Imbens (2001), except for a lighter notation to improve readability and for a change from the letter T to D to be more in line with the rest of the literature. The original assumptions of Imbens (2001), with T replaced by D , are: Monotonicity: For all pairs (z, z') either $D_i(z) \geq D_i(z')$ for all i or vice versa $D_i(z) \leq D_i(z')$ for all i . Exclusion Restriction: $Y_i(Z_i, 0), Y_i(Z_i, 1) \perp\!\!\!\perp Z_i | \{D_i(z)\}_{z \in Z}$. Unconfounded Instrument: $\{Y_i(z, D_i(z))\}_{z \in Z}, \{D_i(z)\}_{z \in Z} \perp\!\!\!\perp Z_i$.

This exclusion restriction requires that the minimum quota does not directly affect the employment prospects of an unemployed person in the case of nonparticipation. In other words, when not participating in ALMP), the employment chances should not depend on the location of residence. This assumption essentially requires the labour market to be fully integrated across the border. Furthermore, any scale or spill over effects due to different treatment incidence on the two sides of the border must be arbitrated away.¹⁰ To take account of these requirements, in Section 4 we consider only local labour markets with very good commuting infrastructure and short commuting times. In particular, we require a commuting distance of at most 30 minutes by car. In addition, the unemployed person also has to be sufficiently mobile to take advantage of this commuting infrastructure to exploit employment opportunities within the entire local labour market. Therefore, we restrict our analysis to unemployed persons without (known or probable) restrictions to their mobility.

2) Regional treatment intention does not affect the treatment outcome: $Y_i^1 \perp\!\!\!\perp Z_i \mid T_i$

This assumption requires not only an integrated labour market, but assumes further that the treatment intention does not affect the quality and composition of ALMP. Essentially, it requires that the quality and composition of the offered programmes (in terms of their impact on employment chances) are the same in both regions. This assumption is more difficult to justify. It appears reasonable that the courses and programmes are of comparable quality on both sides of the border because often they are run by private providers that operate in the whole country.¹¹ However, the quota per unemployed might affect the composition of the labour market programmes. For example, if the quota is low, only one category of programmes might be offered, e.g. training. Once this category is full, an increasing quota induces the REOs to assign unemployed also to other programmes. Such a systematic relationship would imply that the very definition of

¹⁰ A large number of unemployed persons participating in labour market programmes may affect the labour market situation of the non-participants. If the local labour market is integrated, these scale effects (or general equilibrium effects) affect the non-participants on both sides of the border similarly.

¹¹ Some cantons also purchase places from providers in the neighbouring cantons, e.g. Appenzell and StGall.

treatment depends on the quota per unemployed. Such systematic patterns we could not find, though.¹² However, also non-systematic variation in the composition of ALMP, e.g. due to the preferences of the REO management, could invalidate the above assumptions. Some REOs might prefer employment programmes, other rather prefer courses or temporary wage subsidies. Therefore we are going to compare only regions with a similar ALMP structure.

3) Regional treatment intention is independent of employability of local population: $Y_{i,z} \perp\!\!\!\perp Z_i$

Assumption 3 requires that the populations residing on the two sides of the border are identical in terms of their employability. A violation of this assumption could arise due to three channels. (A) Anticipation of the future employment prospects could have affected the instrument. In our application this is not possible because the value of the instrument is computed according to a pre-determined formula where forecasts of labour market developments did not enter. (B) Common factors might have affected the instrument and the outcome variables. Since the quota per unemployed is determined by the population share in 1996, the unemployment share in 1996 and the number of unemployed in 1998, the average employability of the population might have affected the instrument and be related to the employment outcomes measured in 1999, thereby introducing a correlation between the instrument and the potential outcomes.¹³ Nevertheless, such a

¹² Table C.3 gives the composition of the ALMP in the REOs in the local labour markets. It shows the allocation of the treated to training, employment programmes, temporary wage subsidies and other programmes. When regressing the quota per unemployed of Table 2.1 on the average ALMP compositions in the cantons from Table C.3, all coefficients are insignificant (except for the constant). A regression of the quota on the share in training yields a coefficient of 0.12 with a t-statistic of 1.3. A regression of the quota on the share in employment programmes gives -0.05 (t-statistic 0.4), and a regression on the share in temporary wage subsidies gives -0.15 (1.1). When regressing on the training share *and* the employment programme share, the t-statistics are 1.4 and 0.8. Similarly, the t-statistics are 0.8 and 0.5 when training share and temporary wages subsidy share are included as regressors, and 0.6 and 1.1 for employment programme share and temporary wages subsidy share. Only the rather tiny fraction of *other programmes* seems to be negatively related to the quota.

¹³ For example, a low unemployment rate in 1996 and 1998 would have resulted in a rather high quota per unemployed. If the good employment prospects continued into 1999, the instrument and the employment outcomes were positively correlated leading to upward biased estimates. However, other factors could have induced a negative correlation, for example if the higher educated are more likely to migrate between cantons. Because the value of the instrument depends on the population and unemployment *shares* in 1996 divided by the *number* of unemployed in

correlation is unlikely to be large, though, since the quota is based on average figures for the *entire* canton, whereas the IV estimator uses only the observations living close to the cantonal border. While differences in the industrial structures in other parts of the canton can induce differences in the value of the instrument, the composition of the population living *close* to the border is most likely to be very similar on both sides of the border. A potential factor that might have led to a sorting of the population in the vicinity of the border is local taxes, because income tax rates differ between cantons. However, income tax rates also differ between municipalities, which partly off-sets the differences between cantonal tax rates, as municipalities compete for high-earning residents. In addition, differences in tax rates are largely capitalised in the prices for housing and furthermore affect particularly people with high earnings, which represent only a tiny fraction of the unemployed.

(C) Finally, one could argue that the value of the instrument itself might have induced a sorting of the population. Unemployed persons might have realised that the probability of being assigned to labour market programmes is different in the neighbouring canton. As some of them had preferences to take part in programmes, or conversely, to avoid them, they might have preferred attending a REO in the other canton. This, however, would have required moving to the neighbouring canton (before being assigned to a programme). While the costs of changing residence are quite substantial, its benefits are small and highly uncertain. First, the differences in the probability of being assigned to treatment are not very large between neighbouring cantons. Second, as the probability of treatment assignment also depends on the interaction between caseworker and the unemployed, a change in residence could have had unexpected consequences and might have made the unemployed person even worse off. Hence, a sorting of the population induced by the instrument appears unlikely.

4) Regional treatment intention is independent of type-composition of local population: $T_i \perp\!\!\!\perp Z_i$

1998, different cantonal population growth rates could have generated a negative correlation. In-migration would have reduced the quota per unemployed, because of an increase in the *number* of unemployed persons in 1998. Analogously, out-migration would have increased the quota per unemployed. If the higher educated (which enjoy better employment prospects) are over-represented among the migrants, the instrument and the employment outcomes would be negatively correlated.

A further condition is that the value of the instrument should also be independent of the compliance type of the local population. Hence, the portions of always-participants, never-participants, defiers and compliers should be the same on both sides of the border. In particular, this rules out a selective choice of residence of the compliers, whose treatment status depends on the value of the instrument. As discussed above with assumption 3, such a selective choice of residence appears rather unlikely.

5) The size of the defier population is zero: $P(T_i = d) = 0$

This monotonicity assumption requires that an increase in the quota per unemployed does not induce any unemployed person to switch from participation to non-participation. This assumption holds if an increase in the minimum quota would imply that more unemployed are assigned to treatment, but would not lead to any organisational changes. If increases in the quota are also associated with changes in the organisation, some unemployed persons who were assigned to treatment in the regime with the lower quota might no longer be assigned if an increased quota changes the patterns of people assigned to programmes. Although the latter cannot be ruled out, the portion of defiers is likely to be small and thus would not bias the results very much. Furthermore, if the treatment effect for the compliers and the defiers were identical, the bias would even be zero.

6) The size of the complier population is positive: $P(T_i = c) > 0$

The final condition requires that the instrument indeed has an effect on the treatment probability. Table 2.1 already confirmed this. Although this assumption is most innocuous and can even be tested, the size of the complier population does affect substantially the small sample properties of the estimator, as will be discussed below.

Under these conditions, the effect of participation in ALMP for the compliers living in the local labour market along the administrative boundary is identified. It should be noted that with a single instrument at hand and without imposing further structure, only the effect of the active labour market programmes as a

whole is identified, but not the effects of single programmes, e.g. of employment programmes or training courses.¹⁴

3.2 Aggregated LATE

The previous discussion referred to a simple setting with only two neighbouring regions. In many applications, however, there may be multiple administrative regions with multiple borders. The instrument then takes more than two different values, i.e. $Z_i \in \{\underline{z}, \dots, \bar{z}\}$, where $\underline{z} \leq \dots \leq \bar{z}$. If all regions are located in the same local labour market and the above assumptions hold for a multivalued Z , the treatment effect on the subpopulation of *all* compliers is given by:

$$E[Y^1 - Y^0 | T = c^{\underline{z}, \bar{z}}] = \frac{E[Y | Z = \bar{z}] - E[Y | Z = \underline{z}]}{E[D | Z = \bar{z}] - E[D | Z = \underline{z}]}, \quad (2)$$

where $c^{\underline{z}, \bar{z}}$ denotes the population changing treatment status when the value of the instrument changes from \underline{z} to \bar{z} . The treatment effect is identified from the observations in the canton with the highest quota per unemployed and the canton with the lowest quota. In principle, the observations from the cantons with intermediate values of Z are not needed. (This will be explained in more detail below.)

This simple set-up changes, however, when the multiple regions contain various local labour markets with different economic conditions. As the instrument Z may be related to the economic conditions in the region, assumption 3 would be violated if all regions were examined together. This unconfoundedness assumption on Z may be valid only in confined areas along the border between two (or more) regions. Only in these areas along an internal border can a treatment effect be identified. Suppose there are l different local labour markets with an internal border, and for each a treatment effect is estimated. Examining the effects for the l different local labour markets separately can provide useful information on effect heterogeneity. More often, however, this examination of effect heterogeneity may lead to the conclusion that all estimates are in-

¹⁴ One might have thought to isolate the effect of, for example, employment programmes by deleting all observations that participated in any other type of labour market programme. For identification, however, we also would need to

significant and that the differences in the estimates largely reflect noise. To learn anything from the statistical analysis, the information contained in these l different estimates needs to be condensed through aggregation or weighting to obtain some kind of summary measure.

Let $\hat{\theta}$ denote the column vector of the estimated treatment effects for the l local labour markets, and w a vector of weights that sum up to one. Define the *aggregated local average treatment effect* $\hat{\Theta}$ as the weighted average of the single estimates:

$$\hat{\Theta} = w' \hat{\theta}. \quad (3)$$

Since the treatment effects are estimated with different precision in the various local labour markets, we propose to choose weights either to minimise the variance of the aggregated effect,

$$w_{Var}^{opt} = \arg \min_w Var(w' \hat{\theta}) \quad s.t. \quad w' \iota = 1, \quad (4)$$

or to minimise its mean squared error,

$$w_{MSE}^{opt} = \arg \min_w MSE(w' \hat{\theta}) \quad s.t. \quad w' \iota = 1, \quad (5)$$

where ι is a column vector of ones.¹⁵ By the usual considerations, the optimal weights for both cases are given by:

$$w_{Var}^{opt} = \frac{[Var(\hat{\theta})]^{-1} \iota}{\iota' [Var(\hat{\theta})]^{-1} \iota}, \quad w_{MSE}^{opt} = \frac{[MSE(\hat{\theta})]^{-1} \iota}{\iota' [MSE(\hat{\theta})]^{-1} \iota}. \quad (6)$$

The aggregated LATE can be given two interpretations. If the complier treatment effects were identical in all the l local labour markets, $\hat{\Theta}$ is the *minimum-distance estimator* of the common effect. If the treatment effects were heterogeneous, the aggregated LATE gives an approximate estimate of the effect associated

delete all observations that would participate in other types of programmes if the value of their instrument were changed. Since these persons are not identified, this is not possible.

¹⁵ Similar considerations can be found in Heckman and Smith (2000) and Black, Smith, Berger and Noel (2003) on the aggregation of impact estimates from randomized experiments at multiple sites or at different times.

with a change in the scale of the ALMP policy. Although the l different estimates refer to different subpopulations, all these subpopulations have in common that they are at the margin of being assigned to a programme. If the scale of the policy were enlarged or reduced, e.g. through the use of quotas, these individuals are the first to be affected. One possibility would be to weight these l different estimates by the estimated size of the respective complier subpopulations. This would assign weights to the different complier subpopulations in proportion to their share of the total subpopulation of all compliers. Weighting by the empirical complier sizes, however, led to a high variance in the Monte Carlo simulations in Appendix A. Another possibility would be to use theoretical weights that would emanate from specific policy scenarios. E.g. if a uniform increase in the share of participants in ALMP in the whole country were aimed for, the necessary fraction of compliers would be the same in all cantons. This would suggest a uniform weighting of the l different estimates. Similarly, if a uniform share of participants were to be achieved, larger numbers of compliers would be needed in cantons with a low ALMP participation rate. An implicit assumption in these policy simulations is that the treatment effects for the ensuing complier subpopulations are similar to those for the compliers reacting to the minimum quota. These theoretical weights, however, might not be efficient from a statistical perspective because they completely neglect the variance of the estimates and may place disproportionate weight on a few very imprecise estimates. The proposed aggregated LATE with weights given by equation (6), on the other hand, is not bound to a specific policy scenario. Rather, it is based on statistical considerations to discern a main pattern from many noisy estimates.¹⁶

For computing the weights (6), an estimate of the variance or the MSE of $\hat{\theta}$ is needed. If all local labour markets are defined as mutually exclusive, all covariance elements of $\hat{\theta}$ are zero. However, local labour markets might also be defined as overlapping, and this could sometimes lead to more precise estimates in finite samples. Consider a local labour market that extends into three administrative regions: r , s and t . Let z_r , z_s and z_t denote the value of the instrument in the three respective regions and suppose that

¹⁶ One could imagine combining these two approaches, by choosing weights that approximately correspond to a specific policy scenario but also take the precision of the l different estimates into account. This is not attempted in this paper, though.

$z_r < z_s < z_t$. Instead of considering this as a single local labour market, one could think of using the border between r and s , the border between s and t and the border between r and t separately as instruments. The respective complier treatment effects for these three borders are:

$$\theta_{r|s} = \frac{E[Y | Z = z_s] - E[Y | Z = z_r]}{E[D | Z = z_s] - E[D | Z = z_r]} \quad \theta_{s|t} = \frac{E[Y | Z = z_t] - E[Y | Z = z_s]}{E[D | Z = z_t] - E[D | Z = z_s]}$$

and

$$\theta_{r|t} = \frac{E[Y | Z = z_t] - E[Y | Z = z_r]}{E[D | Z = z_t] - E[D | Z = z_r]} \quad (7)$$

When passing the border from region r to region s , the size of the complier subpopulation is $E[D | Z = z_s] - E[D | Z = z_r]$. The size of the complier subpopulation associated with crossing the border from s to region t is $E[D | Z = z_t] - E[D | Z = z_s]$. When passing from r to t , the complier subpopulation is the union of the two previous subpopulations. It can easily be seen that the treatment effect $\theta_{r|t}$ is identical to a weighted average of $\theta_{r|s}$, $\theta_{s|t}$ and also of $\theta_{r|s}$, $\theta_{r|t}$, $\theta_{s|t}$, when the separate effects are weighted by the sizes of the corresponding complier subpopulation (see also Heckman and Vytlačil, 1999).

Hence, the treatment effect for this local labour market can either be identified by $\theta_{r|t}$ and neglecting the observations in region s . Alternatively, it can be obtained by estimating the pair-wise effects $\theta_{r|s}$, $\theta_{r|t}$ and $\theta_{s|t}$ and weighting them by the estimated complier fractions. Instead of weighting these pair-wise effects by the complier fractions, using the variance or MSE-minimising weights (6) could lead to more precise estimates in finite samples. The Monte Carlo simulations in Appendix A show that this can indeed be the case.

Estimation of the aggregated treatment effect proceeds in three steps. First, $\hat{\theta}$ is estimated for all local labour markets. Second, the covariance matrix of $\hat{\theta}$ is estimated. Finally, the optimal weights are computed and $\hat{\Theta} = \hat{w}'\hat{\theta}$ is calculated. The treatment effect for each of the local labour markets can be estimated conveniently by replacing means by averages in (1). Since for each local labour market the instrument is bi-

nary, this Wald estimator is equivalent to the linear IV or the two-stage least squares (2SLS) estimator. However, it is well known that IV estimators often perform poorly if the instrument is weak. Their finite sample distributions may be fat tailed and moments may exist only in the case of overidentification. This latter problem is of particular concern, as the very definition of the optimal weights (6) assumes the existence of the second moment.

3.3 Weak instruments and aggregated LATE

The literature on “weak instruments” has taken various directions to deal with these issues. One strand of the literature intends to improve inference for IV estimators in the presence of weak instruments.¹⁷ A second strand developed tests for weak instruments, e.g. Hahn and Hausman (2002) and Stock and Yogo (2002).¹⁸ A third strand of the literature aimed at developing alternative estimators with better finite-sample properties in the case of weak instruments. Many of these are k -class estimators, introduced by Theil (1958) and Nagar (1959), and are of the type:

$$\{D_N' [I_N(1-k) + kP_N] D_N\}^{-1} D_N' [I_N(1-k) + kP_N] Y_N, \quad (8)$$

where N is the sample size, Y_N is the data vector of the outcome variable, D_N is the data matrix of the endogenous variables (including a constant), I_N the identity matrix and $P_N = Z_N(Z_N'Z_N)^{-1}Z_N'$ the projection matrix of the data matrix of instruments Z_N (including a constant). The constant k defines the specific estimator: $k=1$ gives the conventional 2SLS estimator. $k=0$ corresponds to OLS. Choosing k as the smallest root of the determinantal equation $|\Xi' \Xi - k \Xi'(I_N - P_N) \Xi| = 0$, where $\Xi = [Y_N : D_N]$ is the hori-

¹⁷ Alternative asymptotic theory (*many-instrument asymptotics* in Bekker, 1994, Hahn, 2002; *weak-instrument asymptotics* in Staiger and Stock, 1997, Wang and Zivot, 1998, Zivot, Startz and Nelson, 1998, Stock and Wright, 2000; and *higher-order asymptotics* as in Nagar, 1959, Anderson and Sawa, 1979, Morimune, 1983, Rothenberg, 1983, Hahn, Hausman and Kuersteiner, 2002, among others) has been developed, which leads to better coverage probabilities in the presence of weak instruments.

¹⁸ Donald and Newey (2001) are concerned with choosing the appropriate subset out of a set of valid instruments to minimise mean squared error. If many instruments are available, including all instruments in the IV estimator may

zontal concatenation of Y_N and D_N , gives the Limited Information Maximum Likelihood estimator (LIML). In the case of exact identification, LIML and 2SLS are identical. For normal errors (or errors with even fatter tails), LIML does not possess finite moments and 2SLS has moments only in the case of overidentification.¹⁹ Fuller (1977) proposed a modified LIML estimator with $k = k_{LIML} - \alpha / (N - L)$, where α is a positive constant and L is the number of instruments. The Fuller estimator has first and second moments in finite-samples. Choosing $\alpha = 1$ gives nearly unbiased estimates, whereas $\alpha = 4$ yields asymptotically the smallest MSE, see also Rothenberg (1984). Angrist, Imbens and Krueger (1999), Blomquist and Dahlberg (1999), Donald and Newey (2001), and Hahn, Hausman and Kuersteiner (2002) proposed alternative estimators, which have superior theoretical properties compared to 2SLS and often perform well in finite-samples, see also Stock, Wright and Yogo (2002).

The finite-sample behaviour of these estimators, however, has been analysed so far only under conditions very different from our setting. Often designs with a large number of overidentifying continuous instruments were considered, whereas our instrument is binary. In addition, we estimate not only parameters but also the optimal weights for averaging these parameters. Since this set-up is rather different from conventional instrumental variables estimation, we conducted a Monte Carlo study to analyse the properties of the aggregated treatment effect estimator in Appendix A. Although the limited Monte Carlo study does not allow for very strong conclusions, in total, the Fuller (1) estimator with variance-weighting performed best.

worsen finite-sample properties when some of the instruments are weak. The data-driven procedure of Donald and Newey to select a subset of the available instruments can improve upon finite-sample properties.

¹⁹ See Basmann (1961, 1963), Kabe (1964), Richardson (1968), Sawa (1969), Mariano and Sawa (1972), Nelson and Startz (1990), Buse (1992), Maddala and Jeong (1992) and others.

4 Implementation of the evaluation of Swiss active labour market policies

4.1 Data and sample selection

The population for the microeconomic evaluation are all individuals who were unemployed on the first of January 1998, for at most one year.²⁰ For a random sample of 81'399 individuals, detailed information on employment histories (including self-employment), monthly earnings, participation in ALMP and personal characteristics for the years 1988 to 1999 were obtained from administrative databases of the unemployment insurance system and the social security records. The information includes age, gender, marital status, household composition, place of residence, nationality, type of work permit, mother tongue, foreign languages, education, qualification, caseworker's assessment of employability, position in last job, occupation and industry of last job, size of town where worked before, looking for part-time or full-time job, occupation and industry of desired job, information on earnings in last job, duration of contribution to unemployment insurance, disability etc. For more details see Appendix B.

Several sample selection criteria are applied to restrict the population to individuals who are *eligible* to take part in ALMP and for whom no restrictions to their mobility are known or probable, as discussed with our IV identification strategy in Section 3.1. In particular, disabled persons are excluded, as well as foreigners with a working permit of less than a year, since there are legal restrictions to their mobility. In addition, persons with a very low income (monthly earnings in last job below 1000 CHF, ≈650 EURO) are left out, because monetary costs of commuting might be an obstacle to them to take advantage of job opportunities that are not nearby. We also restrict the sample to the prime age group (25-55). Furthermore, we excluded students, apprentices and homeworkers and persons who are registered as part-time employed. The losses of observations due to these sample selection rules are detailed in Table B.1 in the appendix. The final sample contains 66'713 observations.

²⁰ Persons who were unemployed for more than a year are excluded because they entered in unemployment before the reform was enforced in January 1997 and were thus subject to different rules and regulations at the entry in unemployment. In addition, their entitlement period to ALMP expired before the end of 1998.

4.2 Variables

Since with a binary instrumental variable we can identify only the effect of a binary treatment, as discussed in Section 3, we partition the population into two groups. Treatment is defined as entering in an active labour market programme of at least one week duration at any time in 1998. These programmes vary in quality and content but can broadly be distinguished into 4 categories: *training*, *employment programmes*, *temporary wage subsidies*, and *other* programmes.

Table 4.1: Descriptive statistics of selected characteristics (Means or shares)

Variable name		Treated	Non-treated
Age		38.2	37.3
Female		45	41
Swiss nationality		56	55
Insured earnings (Swiss Francs, CHF) ^a		3'980	3'832
<u>Treatment status in 1998 (first programme)</u>			
Participated in training		34	-
Participated in employment programme		22	-
Participated in temporary wage subsidy		42	-
Participated in other labour market programmes		1	-
<u>Outcome variables in 1999</u>			
Employed	average for January-April 1999	45	53
	average for May-August 1999	59	63
	average for September-December 1999	66	64
High earnings (> 90% of insured earnings)	average for January-April 1999	27	35
	average for May-August 1999	35	42
	average for September-December 1999	37	42
Labour earnings (from wage and self-employment) in CHF	average for January-April 1999	1'714	2'091
	average for May-August 1999	2'218	2'455
	average for September-December 1999	2'408	2'470
Observations		40'193	26'520

Note: 66'713 observations. ^{a)} 1 Swiss Franc (CHF) \approx 2/3 Euro.

Overall, 60% of all unemployed entered active labour market programmes (with the average month of entry being March 1998). Table 4.1 shows that 34% of the treated received training, 22% participated in an employment programme, 42% became employed in a temporary wage subsidy, and the remaining 1% participated in other labour market programmes.²¹

²¹ The numbers refer to the first programme the unemployed participated in. There are some unemployed participating in more than one programme, however most subsequent programmes are of the same type as the first programme.

To uncover the dynamics of the programme effects, we follow the individual labour market situation over the year 1999 and create the following outcome variables: A person is defined as being *employed* if positive earnings are reported for at least three consecutive months. By this definition we intend to measure continuous employment of at least *one month* duration.²² A person is being considered as *employed with high earnings* if she earns at least 90% of what she earned on her last job.²³ In addition, *monthly earnings* itself is considered as a further outcome variable. These three outcome variables are created for every person and every month in 1999. To reduce noise, 4-month averages are taken. The last panel in Table 4.1 shows their sample means. These outcome variables capture the different policy objectives of the active labour market policies, which are rapid and lasting re-employment without large earnings losses.

4.3 Identifying local labour markets

To apply the evaluation strategy discussed in the previous sections, integrated *local labour markets* with internal administrative borders need to be found. We define a local labour market in terms of the area corresponding to one or more regional employment offices. In particular, we seek to identify a *cluster of REOs* that satisfies the following criteria: 1) The REOs belonging to the cluster are spread over 2 cantons; 2) commuting times by car between these REOs are 30 minutes or shorter; 3) the same language (French, German or Italian) is spoken in the areas belonging to the REOs; 4) The ALMP composition is similar in the REOs. With the first criterion, we identify local labour markets pair-wise between cantons. As discussed in Section 3.2, if a local labour market extends into three or more cantons, we consider each pair-wise comparison between the cantons separately.

The second criterion ensures that all potential employers can be reached within convenient commuting distance from both sides of the cantonal border. This criterion is implemented by examining the distances be-

²² If positive earnings are reported in only two consecutive months, this could, for example, be the result of an employment spell that lasted only one week and began close to the end of the first month. By requiring three consecutive months, such very short spells can largely be excluded. However, some short employment spells might still be included if they appear repeatedly within this three months window.

tween any pair of regional employment offices in terms of commuting times by car. (See Table B.3 for a list of REOs. The commuting times are shown in Table C.1. A maximum driving time of about 30 minutes appears to be acceptable for exploiting wage arbitrage opportunities.²⁴ Switzerland is one of the countries with the highest per capita car ownership world-wide. In addition, public transportation is also very good in Switzerland and reaches every village.

The third criterion takes account of the different language regions, as Switzerland consists of German, French and Italian speaking parts. Local labour markets where French is spoken on the one side of the border and German on the other side are excluded. French-German bilingual regions bordering to German speaking regions are not excluded, though. In such local labour markets, all observations with French mother tongue are deleted, as they may not consider the neighbouring German speaking region as part of their labour market where to search for jobs. According to the criteria one to three, 30 local labour markets are identified (and shown in Table C.2).

As discussed in Section 3, one of the IV assumptions is that the quality and type of treatment is identical on both sides of the cantonal border. It appears reasonable to assume that the quality of the services does not vary systematically between neighbouring regions, because many of the providers of training courses and employment programmes operate in the whole country. However, there seems to be variation in the types of programmes the caseworkers assign their clients to, which may be due to different philosophies of the REOs. Criterion 4 therefore requires that the allocation of the treated to the different ALMP categories is similar on both sides of the border. By defining similarity in terms of *maximum divergence* between the allocations, we distinguish between local labour markets with a *very similar* ALMP-structure (Selection A), labour markets with a *comparable* structure (Selection B, including Selection A), and labour markets with a

²³ The social security records data report earnings for every month. Unlike in the U.S., the social security earnings data are *not* top-coded. In addition, every regularly employed person has to contribute, including government workers and self-employed.

²⁴ In Switzerland about 50% of the working population commute less than 15 minutes, about 80% commute 30 minutes or less (one way; Bundesamt für Statistik, 2003).

distinct ALMP structure. The implementation of this similarity criterion is described in Appendix C, and the Tables C.2 and C.3 give details on all 30 local labour markets.

Table 4.2 displays summary statistics only for the local labour markets with at least a comparable ALMP-structure (Selection B), because only these local labour markets are used for the estimation of the aggregated LATE. The last column indicates whether the labour market is classified as having a very similar ALMP-structure (Selection A). Column one, gives the number of the local labour market, as used in Table C.2. Column two indicates the cantonal border that functions as the instrument, and columns three and four give the REOs belonging to this local labour market (on the two sides of the border). For example, the row corresponding to labour market 7 indicates that the areas corresponding to the REOs of Langenthal and Zofingen are considered as a local labour market, with Langenthal belonging to the canton Bern (BE) and Zofingen belonging to the canton Aargau (AG). Columns five and six give the number of observations in the sample, 313 in Langenthal and 472 in Zofingen. Columns seven and eight display how many of these observations were treated, and column nine gives the difference in the treatment probability. This is an estimate of the fraction of compliers and lies in the range of ± 15 percentage points, with many small values.

Table 4.2: Local labour markets divided by administrative border

#	Cantons	Regional employment offices		Number of observations		% Treated		Complier ^b	Diff. in instrument Table 2.1 ^c	
				n ₁	n ₂	(7)	(8)		(10)	A
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2	SO-BE	Solothurn, Oensingen, Biberist, Zuchwil	Wangen, Langenthal, Burgdorf	877	818	68	63	5.6	-3.9	x
7	BE-AG	Langenthal	Zofingen	313	472	64	68	-4.9	-1.2	
8	BE-FR	Gümligen, Zollikofen, Köniz, Bern (2x)	Murten, Tafers, Fribourg	2'660 ^a	763 ^a	65	67	-1.7	-0.2	x
9	FR-VD	ChatelSt.Denis	Oron la Ville	107	107	74	59	15.0	3.0	x
10	FR-VD	Romont, Estavayer	Payerne, Moudon	371	355	64	60	3.3	3.0	
11	VD-GE	Nyon	Genf (6x)	576	5'700	57	50	6.5	0.8	x
12	VD-VS	Vevey, Aigle, Montreux	Monthey (2x)	1'580	609	59	66	-6.7	-0.7	
13	BL-BS	Pratteln, Münchenstein, Binningen	Basel (3x)	934	2'081	67	53	14.2	2.1	
15	LU-NWOW	Luzern, Emmen, Emmenbrücke, Kriens	Hergiswil (2x)	1'607	265	64	62	2.4	-10.4	
16	LU-ZG	Luzern, Emmen, Emmenbrücke, Kriens	Zug	1'607	571	64	64	-0.2	-1.5	
17	SZ-UR	Goldau	Altdorf	337	150	69	61	8.8	-1.3	
19	AG-ZH	Baden, Wettingen, Wohlen	Opfikon, Effretikon, Uster, Wetzikon, Bülach, Dietikon, Regensdorf	1'529	4'165	64	58	6.6	3.5	
21	ZH-TG	Winterthur	Frauenfeld	1'221	537	59	69	-9.9	-1.8	x
22	ZH-SG	Meilen, Thalwil	Rapperswil	1'421	360	56	60	-3.8	-4.5	x
23	ZH-SZ	Meilen, Thalwil	Lachen	1'421	529	56	72	-15.2	-8.5	x
24	TG-SH	Frauenfeld	Schaffhausen	537	605	69	63	6.3	3.1	
25	TG-SG	Amriswil	Rohrschach, Oberuzwil	474	853	64	66	-1.5	-2.7	x
28	SG-SZ	Rapperswil	Lachen	360	529	60	72	-11.4	-3.9	

Note: ^a Number of observations after deleting individuals with French mother-tongue, because a French-German bilingual region is bordering a German speaking region.

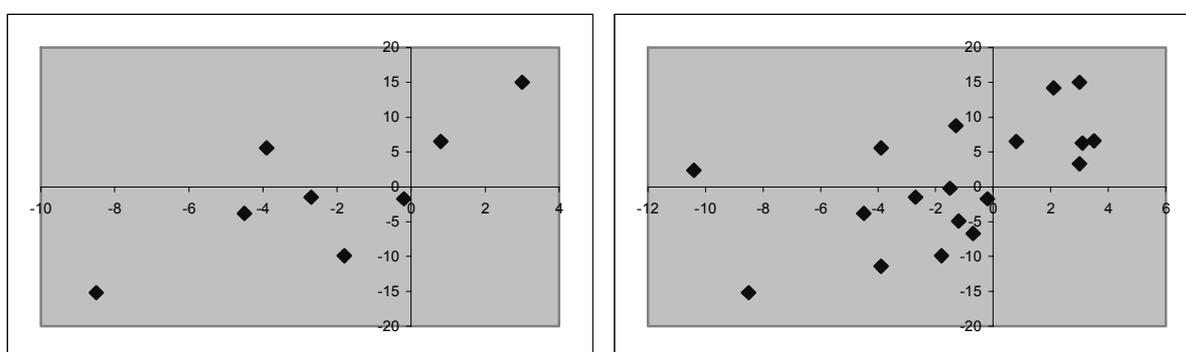
^b The estimate of the fraction of compliers is the difference between the previous two columns.

^c Difference in the instrument quota per unemployed (Table 2.1, column 5) between the two cantons.

Before estimating any treatment effects, it is interesting to compare the variation in actual treatment, as estimated with this micro data set, with the variation in the instrument *quota per unemployed*, as discussed in Section 2. Column ten of Table 4.2 gives the difference in the quota per unemployed between the two cantons (calculated from column 5 of Table 2.1). Column nine shows the estimated fraction of compliers, i.e. the fraction of unemployed who are induced to treatment when being subjected to the quota of the other canton. This is the effect of the instrument Z on the probability of treatment. Although not strictly necessary for identification, since identification in this paper is only based on pair-wise comparisons, a positive relationship between differences in the instrument Z and differences in the treatment probability is expected. Figure 4.1 shows that this relationship is indeed positive (on the left for Selection A, on the right for Selec-

tion B). The correlation is 0.78 for Selection A and 0.57 for Selection B. In all labour markets, the change in Z and its estimated effect on treatment participation D have the same sign, except for the three labour markets 2, 15 and 17. The opposing signs in these latter markets may be the result of sampling variability. On the other hand, it might also indicate a failure of the monotonicity condition in these three markets. To be on the safe side, these three local labour markets are excluded in the subsequent estimations. (The results are similar when including them.)

Figure 4.1: Correlation between differences in the instrument and the estimated complier fraction



Note: Abscissa: Differences in the quota per unemployed (column 10 of Table 4.2); ordinate: differences in treatment probability (complier fraction, column 9 of Table 4.2); for labour markets of Selection A (left figure) and of Selection B (right figure).

5 Results

5.1 Local average treatment effects in the local labour markets

In the following the estimation results are presented. Table 5.1 shows the Fuller estimates $\hat{\theta}$ of the local average treatment effects for all 30 labour markets. The corresponding Wald estimates are given in the appendix in Table D.1. The effects of being treated in 1998 on the employment and earnings outcomes in 1999 are given, separately for the months May to August 1999 and September to December 1999. The estimation results are highly volatile and sometimes even change sign between the different outcome variables. Some of the estimates are even outside the logical range, e.g. a reduction in employment probability of more than 100 percentage points. This volatility may indicate a substantial heterogeneity in the treatment effects. More plausible, though, it is the result of substantial noise in the estimates, reflecting the weakness of the instrument Z and the often small number of observations.

Table 5.1: Fuller estimates of the local average treatment effects for the 30 local labour markets

Region	LATE estimated with Fuller (1), year 1999						
	Employment ^a (in%-points)		High earnings ^b (in%-points)		Labour earnings in CHF ^c		
	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.	
1	SO-BE	1	-12	-25	-35	740	110
2	SO-BE	-143	-97	-67	-43	-4'650	-2'800
3	SO-BL	12	25	-3	-9	-460	-260
4	SO-AG	-5	-37	-12	-28	-2'320	-3'970
5	NE-BE	41	55	58	74	1'610	790
6	NE-BE	0	-1	7	13	410	-370
7	BE-AG	-176	-127	-44	-29	-3'430	-2'290
8	BE-FR	-110	-63	1	27	-2'550	-1'390
9	FR-VD	79	38	56	40	3'020	1'720
10	FR-VD	22	54	51	62	3'320	4'760
11	VD-GE	97	147	56	79	7'560	9'110
12	VD-VS	78	77	47	34	1'220	950
13	BL-BS	-5	5	21	10	1'620	1'670
14	LU-AG	-121	-90	-104	-96	-5'260	-4'840
15	LU-NWOW	-126	-149	-74	-114	-7'390	-9'240
16	LU-ZG	-17	-8	-9	-3	350	600
17	SZ-UR	-17	-62	13	-3	-1'270	-3'720
18	SZ-ZG	6	-25	-17	-48	-6'250	-7'550
19	AG-ZH	16	16	54	67	-250	-100
20	ZH-SH	-32	-58	-21	-41	-4'820	-5'750
21	ZH-TG	35	16	2	-10	190	-440
22	ZH-SG	3	-65	0	-48	-4'810	-6'880
23	ZH-SZ	16	9	1	-2	-660	-990
24	TG-SH	54	66	12	28	2'190	2'200
25	TG-SG	-5	-24	-8	-41	1'280	-160
26	SG-AR	14	-3	16	26	3'120	4'180
27	AR-AI	-7	-11	-8	-16	-810	-1'110
28	SG-SZ	18	44	0	20	1'490	2'130
29	SG-GR	-74	-88	-40	-21	-4'460	-3'200
30	GR-TI	-104	-9	-68	-10	-4'680	-1'430

Note: ^a Employment: A person is employed if positive earnings are reported for at least 3 consecutive months.
^b High earnings: A person is being considered as employed with high earnings if reported earnings are larger than 90% of the insured earnings.
^c Earnings: From wage and self-employment (excluding benefits). 1 CHF \approx 2/3 Euro.

5.2 Aggregated LATE

To learn anything from these noisy estimates, their statistical information is condensed to an average effect of ALMP for compliers. The estimated effects of Table 5.1 are aggregated, using variance weighting as suggested by the Monte Carlo study in Appendix A. The weights are $w = \hat{\Sigma}^{-1} \mathbf{1} / (\mathbf{1}' \hat{\Sigma}^{-1} \mathbf{1})$ where $\hat{\Sigma} = \text{Var}(\hat{\theta})$ is the estimated covariance matrix of $\hat{\theta}$. Table 5.2 shows the aggregated treatment effect $\hat{\Theta} = w' \hat{\theta}$ for the 15 local labour markets with comparable ALMP-structure (Selection B, without labour markets 2, 15, 17).

The first rows in Table 5.2 give the estimated weights, which were estimated separately for each outcome variable. The local labour market 13 receives about 30% of the weights, while the remaining weights are spread relatively even among the other labour markets. The following three rows show the estimated aggregate effect $\hat{\Theta} = w' \hat{\theta}$, the estimated standard error $\sqrt{w' \hat{\Sigma} w}$, and the t-value $w' \hat{\theta} / \sqrt{w' \hat{\Sigma} w}$. The covariance matrix $\hat{\Sigma}$ is estimated through 10'000 bootstrap replications.

For testing whether $\hat{\Theta}$ is statistically different from zero, two approaches are pursued. In the first approach, the estimated weights \hat{w} are treated as non-stochastic and the distribution of the aggregated effect $\hat{\Theta}$ is simulated through 10'000 bootstrap replications. The quantiles of the bootstrap distribution (centred by subtracting the estimate $\hat{\Theta}$) are shown in the rows labelled '*Bootstrap of estimate with fixed weights*'. The largest quantile that is surpassed by the estimate $\hat{\Theta}$ is underlined.

In the second approach, the weights \hat{w} are themselves treated as stochastic in the simulation of the distribution of the aggregated effect $\hat{\Theta}$. This, however, requires a double bootstrapping approach. First, a bootstrap sample is drawn and $\hat{\theta}_b$ is estimated from this bootstrap sample. The estimation of the weights \hat{w}_b for this bootstrap sample requires an estimate of the covariance matrix of $\hat{\theta}_b$. This covariance matrix $\hat{\Sigma}_b = \text{Var}(\hat{\theta}_b)$ is estimated by bootstrapping from the bootstrap sample. Only 1'000 replications are used to estimate $\hat{\Sigma}_b$, since the whole process is computationally demanding. With $\hat{\Sigma}_b$ estimated, the weights are calculated as $\hat{w}_b = \hat{\Sigma}_b^{-1} \iota / (\iota' \hat{\Sigma}_b^{-1} \iota)$ and the bootstrap estimate $\hat{\Theta}_b = \hat{w}_b' \hat{\theta}_b$ is computed. This process is repeated 1'000 times. The quantiles of the distribution of $\hat{\Theta}_b$ (centred by subtracting $\hat{\Theta}$) are shown in the rows labelled '*Bootstrap of estimate with stochastic weights*'. As an alternative to bootstrapping $\hat{\Theta}$, the results for bootstrapping the t-values are also shown.²⁵

²⁵ If the estimated aggregated effect is asymptotically normal, the t-statistic is asymptotically pivotal and a test based on bootstrapping the t-statistic has better properties. Deriving conditions under which this is the case is beyond the scope of this paper. Notice, however, that if the weights are treated as stochastic, the statistic $\hat{w}' \hat{\theta} / \sqrt{\hat{w}' \hat{\Sigma} \hat{w}}$, which

Table 5.2: Aggregated treatment effect for compliers, Selection B

weights (%)	Employment (in%-points)		Outcome variables, year 1999 High earnings (in%-points)		Labour earnings in CHF	
	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.
BE-AG	1	1	3	4	3	3
BE-FR	1	2	3	3	3	3
FR-VD	3	4	4	4	5	5
FR-VD	3	2	2	2	2	1
VD-GE	4	3	5	4	2	2
VD-VS	4	5	5	5	8	9
BL-BS	34	35	30	32	32	31
LU-ZG	1	2	3	3	1	1
AG-ZH	11	12	8	8	12	11
ZH-TG	7	9	8	9	9	9
ZH-SG	7	2	7	3	2	1
ZH-SZ	11	19	10	16	16	16
TG-SH	1	0	1	1	1	1
TG-SG	3	3	3	2	3	4
SG-SZ	9	3	9	4	4	3
Estimate of mean	14.3	13.7	16.8	13.4	702	541
Std.error	7.3	7.2	7.3	7.6	401	396
t-value	1.95	1.90	2.30	1.78	1.75	1.36
Bootstrap of estimate with fixed weights						
Q _{0.95}	<u>13.2</u>	<u>13.7</u>	12.4	<u>13.3</u>	712	716
Q _{0.975}	15.4	16.2	<u>14.8</u>	16.1	855	853
Q _{0.995}	20.2	20.5	20.0	20.6	1'158	1'138
Bootstrap of estimate with stochastic weights						
Q _{0.95}	8.5	9.3	8.2	9.0	422	<u>524</u>
Q _{0.975}	11.4	11.0	10.0	<u>11.1</u>	546	<u>650</u>
Q _{0.995}	<u>13.4</u>	<u>13.3</u>	<u>13.9</u>	14.7	<u>694</u>	813
Bootstrap of t-value with fixed weights						
Q _{0.95}	1.56	1.67	1.44	1.50	1.49	1.50
Q _{0.975}	<u>1.78</u>	<u>1.90</u>	1.63	<u>1.72</u>	<u>1.71</u>	1.71
Q _{0.995}	2.26	2.46	<u>2.11</u>	2.17	2.25	2.12
Bootstrap of t-value with stochastic weights						
Q _{0.95}	1.17	1.30	1.06	1.23	1.10	<u>1.34</u>
Q _{0.975}	1.51	1.55	1.34	<u>1.43</u>	1.35	1.61
Q _{0.995}	<u>1.80</u>	<u>1.82</u>	<u>1.75</u>	1.92	<u>1.67</u>	1.96

Note: Estimation of stderror based on 10'000 bootstrap replications. Quantiles of estimate with fixed weights based on 10'000 replications. The other quantiles are estimated by double bootstrapping and only based on 1'000 replications in inner bootstrap and 1'000 replications in outer bootstrap. Aggregated LATE for labour markets 7, 8, 9, 10, 11, 12, 13, 16, 19, 21, 22, 23, 24, 25, 28.

The estimation results indicate a positive effect of participation in ALMP on employment and high earnings employment of about 15 percentage-points. With the weights treated as stochastic, all estimates are signifi-

is referred to as the t-value in Table 5.2, is not the usual t-statistic because the denominator $\sqrt{\hat{w}'\hat{\Sigma}\hat{w}}$ is not a consistent estimator of the standard-deviation of $\hat{w}'\hat{\theta}$. Therefore, it might be non-pivotal even if the estimated aggregated effect were asymptotically normal. Bootstrapping the t-value requires double bootstrapping, even when the weights are treated as fixed.

cant at the 1% or 5% level, except for the last one, which is significant only at the 10% level. The estimated earnings effect is about 600 CHF. With average previous earnings of about 4'000 CHF (Table 4.1), this is the expected magnitude if the additional 15 percent employed were earning their average previous wage. The estimation results for Selection A (see Table D.2) are largely similar but with smaller and insignificant earnings effects, which may be due to the smaller number of observations included.

5.3 Aggregated treatment effects for urban labour markets

The previous analysis indicated positive effects of labour market programmes on employment prospects. These results, however, might be difficult to interpret as they refer to a heterogeneous population in different types of labour markets. To analyse the effects for a more homogeneous population, Table 5.3 gives results for urban labour markets only.

Table 5.3: Aggregated treatment effect for large-city regions

	Employment (in %-points)		Outcome variables, year 1999		Labour earnings in CHF	
	May-Aug.	Sept.-Dec.	High earnings (in %-points)	Sept.-Dec.	May-Aug.	Sept.-Dec.
weights (%)						
BE-FR	1	2	4	4	3	4
VD-GE	5	3	6	5	2	2
BL-BS	42	41	38	41	41	40
AG-ZH	14	14	11	10	15	15
ZH-TG	10	11	12	12	13	13
ZH-SG	2	1	3	2	1	1
ZH-SZ	27	28	26	27	25	25
Estimate of mean	11.2	10.7	17.8	13.5	525	429
Std.error	8.1	7.9	8.3	8.5	458	459
t-value	1.38	1.36	2.15	1.58	1.15	0.94
Bootstrap of estimate with fixed weights						
Q _{0.95}	14.6	14.4	14.1	14.2	829	809
Q _{0.975}	17.4	16.9	<u>17.2</u>	17.2	973	958
Q _{0.995}	22.5	22.1	22.6	23.0	1'298	1'253
Bootstrap of estimate with stochastic weights						
Q _{0.95}	11.4	11.2	9.8	10.4	573	596
Q _{0.975}	13.9	13.0	11.6	<u>12.8</u>	716	684
Q _{0.995}	17.6	17.2	<u>16.3</u>	15.4	929	945
Bootstrap of t-value with fixed weights						
Q _{0.95}	1.53	1.57	1.46	<u>1.46</u>	1.56	1.50
Q _{0.975}	1.78	1.91	1.67	1.64	1.78	1.75
Q _{0.995}	2.41	2.44	<u>2.12</u>	2.02	2.34	2.16
Bootstrap of t-value with stochastic weights						
Q _{0.95}	1.38	1.41	1.13	1.17	1.29	1.30
Q _{0.975}	1.70	1.66	1.33	<u>1.42</u>	1.54	1.49
Q _{0.995}	2.11	2.16	<u>1.83</u>	1.85	2.00	1.96

Note: See note below Table 5.2. Aggregated LATE for labour markets 8, 11, 13, 19, 21, 22, 23.

As urban labour markets are considered the region between Bern and Fribourg (labour market 8), the region around Geneva (labour market 11), the region around Basle (13), the region around Zurich (19), the region around Winterthur (21) and the south-eastern shores of lake Zurich (22 and 23). See Table C.4 for details. The estimates for these urban regions are slightly smaller but almost the same as in Table 5.2, albeit less significant.

6 Conclusions

In this paper, a new approach to exploiting regional administrative borders locally as an instrumental variable to evaluate the effects of policy interventions has been suggested and applied to the evaluation of Swiss active labour market policies. The methodological contributions are twofold. First, the choice of an appropriate estimator for estimating treatment effects with a *binary* instrument has been examined. It was found that the Wald estimator, which is used in almost all applications with a binary instrument, performs poorly when the instrument is weak. The Fuller (1977) estimator turned out to be most reliable in finite samples. Second, with many different regional borders, separate effects can be estimated for each of these. However, most of them may be insignificant due to the small number of observations in the local areas neighbouring the border. To uncover any significant pattern from these noisy estimates, a weighting scheme for the aggregation of these estimates into a single, specific local average treatment effect has been proposed. The Fuller estimator again seemed to be particularly suited for the estimation of the aggregated treatment effect, because its finite-sample moments exist.

For evaluating the Swiss active labour market policies, the geographical variation in treatment intensity, induced by the 1997 reform of the unemployment insurance system, has been exploited to estimate the effects of participating in labour market programmes on subsequent employment and earnings prospects. Fairly large positive and significant aggregate effects on individual employment probabilities of about 15 percentage-points were found. In addition, no earnings losses were observed, indicating that the higher employment probability was not the result of pushing unemployed into low-paying jobs.

On a first sight, these positive employment effects appear to be at odds with the often negative or insignificant effects found in the literature, e.g. in Gerfin and Lechner (2002), Gerfin, Lechner and Steiger (2002), or Lalive, van Ours and Zweimüller (2000). One difference to these other studies is the aggregation of the active labour market programmes. Whereas many of the other studies distinguish the effects for different types of programmes, in this paper all labour market programmes are aggregated into one group, because disaggregated effects by programme type are not identified with this instrumental variable strategy. Positive treatment effects indeed have also been detected in the other studies for certain programme types. Gerfin and Lechner (2002) and Gerfin, Lechner and Steiger (2002), for example, found large and significant positive effects for the temporary wage subsidies. In the urban local labour markets considered in this paper, temporary wage subsidies represent about 40-50% of the active labour market programmes, while employment programmes account only for about 20-25%.

A second difference is the geographical focus. Whereas the other studies examine average effects for the entire country, in this paper we have to restrict our focus on the individuals living close to a cantonal border. A third difference is that our estimates refer to a (weighted) treatment effect on the *compliers*, whereas most other studies concentrate on population average treatment effects or effects on the treated. The subpopulation of compliers consists of those unemployed that were assigned to active labour market programmes because of differences in the (indirect) incentives to caseworkers to fill the minimum quota. A simple argument would postulate that the treatment effect for this marginal group should be smaller than the treatment effect on the treated if utility maximising individuals decide about participation or if caseworkers act on their behalf. In the case of Switzerland, this need not be the case. With the activation principle introduced by the reform, it is the caseworker who can push an unemployed person into a labour market programme. Lechner and Smith (2003) analysed this assignment process and found that caseworkers were not very successful in assigning unemployed to their most beneficial programmes. Overall, they found that a random allocation presumably would not have been worse. The marginal group of compliers might therefore have benefited more from labour market programmes than others (at least in the short term as measured in this paper, when the economy was booming), and the local average treatment effect thus could have

been higher than the treatment effect on the treated. Summarising this discussion succinctly, the Swiss active labour market programmes seem to have been effective for at least a part of the population, while they may not have been so for the population as a whole. This individual treatment effect heterogeneity may support the need for a better targeting of active labour market programmes.

References

- Anderson, T.W. and T. Sawa (1979): "Evaluation of the Distribution Function of the Two Stage Least Squares Estimator," *Econometrica*, 47, 163-182.
- Angrist, J.D., G.W. Imbens and A.B. Krueger (1999): "Jack-knife Instrumental Variables Estimation," *Journal of Applied Econometrics*, 14, 57-67.
- Angrist, J.D., G.W. Imbens and D.B. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 91, 444-472, with discussion by J.J. Heckman, R.A. Moffitt, J.M. Robins and S. Greenland, and R.P. Rosenbaum.
- Basman, R.L. (1961): "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in Two Leading Overidentified Cases," *Journal of the American Statistical Association*, 56, 619-636.
- Basman, R.L. (1963): "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three-Equation Case," *Journal of the American Statistical Association*, 58, 161-171.
- Bekker, P.A. (1994): "Alternative Approximations to the Distribution of Instrumental Variables Estimators," *Econometrica*, 62, 657-681.
- Black, D., J. Smith, M. Berger and B. Noel (2003): "Is the Threat of Reemployment Services More Effective Than the Services Themselves? - Evidence from Random Assignment in the UI System", *American Economic Review*, 93, 1313-1327.
- Black, S. (1999): "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, 114, 577-599.
- Blomquist, S., and M. Dahlberg (1999): "Small Sample Properties of LIML and Jack-knife IV Estimators: Experiments with Weak Instruments," *Journal of Applied Econometrics*, 14, 69-88.

- Bound, J., Jaeger, D.A., and R.B. Baker (1995): "Problems With Instrumental Variable Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443-450.
- Buse, A. (1992): "The Bias of Instrumental Variable Estimators," *Econometrica*, 60, 173-180.
- Bundesamt für Statistik (2003), *Volkszählung 2000, Pendelverkehr - Neue Definition der Agglomerationen*, Pressemitteilung vom 15.5.2003, Neuchatel.
- Card, D., and A. Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84, 772-793.
- Donald, S.G., and W.K. Newey (2001): "Choosing the Number of Instruments," *Econometrica*, 69, 1161-1191.
- Fuller, W. (1977): "Some Properties of a Modification of the Limited Information Estimator," *Econometrica*, 45, 939-953.
- Gerfin, M., and M. Lechner (2002): "Microeconomic Evaluation of the Active Labour Market Policy in Switzerland," *The Economic Journal*, 112 (482), 854-893.
- Gerfin, M., M. Lechner and H. Steiger (2002): "Does subsidised temporary employment get the unemployed back to work? An econometric analysis of two different schemes," CEPR Discussion Paper No. 3669.
- Hahn, J. (2002): "Optimal Inference with Many Instruments," *Econometric Theory*, 18, 140-168.
- Hahn, J., and J. Hausman (2002): "A New Specification Test for the Validity of Instrumental Variables," *Econometrica*, 70, 163-189.
- Hahn, J., J. Hausman and G. Kuersteiner (2002): "Estimation with Weak Instruments: Accuracy of Higher Order Bias and MSE Approximations," *mimeo Brown University and MIT*.
- Heckman, J., and J. Smith (2000): "The Sensitivity of Experimental Impact Estimates: Evidence from the National JTPA Study," in D. Blanchflower and R. Freeman (eds.), *Youth Employment and Joblessness in Advanced Countries*, Chicago: University of Chicago Press for NBER, 331-356.
- Heckman, J.J., and E.J. Vytlacil (1999): "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, 96:4730-4734.
- Holmes, T. (1998): "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106, 667-705.

- Imbens, G.W. (2001): "Some Remarks on Instrumental Variables", in M. Lechner and F. Pfeiffer (eds.), *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica/Springer, 17-42.
- Imbens, G.W., and J.D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467-475.
- Kabe, D.G. (1964): "On the Exact Distributions of the GCL Estimators in a Leading Three Equation Case," *Journal of the American Statistical Association*, 59, 881-894.
- Lalive, R., J.C. van Ours and J. Zweimüller (2000): "The Impact of Active Labor Market Programs on the Duration of Unemployment," IEW Working Paper No. 41, University of Zurich.
- Lechner, M., and J.A. Smith (2003): "What is the Value Added by Caseworkers?," CEPR Discussion Paper No. 3825.
- Maddala, G.S., and J. Jeong (1992): "On the Exact Small Sample Distribution of the Instrumental Variable Estimator," *Econometrica*, 60, 181-183.
- Mariano, R.S., and T. Sawa (1972): "The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," *Journal of the American Statistical Association*, 67, 159-163.
- Morimune, K. (1983): "Approximate Distributions of k-Class Estimators when the Degree of Overidentifiability is Large Compared with the Sample Size," *Econometrica*, 51, 821-841.
- Nagar, A.L. (1959): "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," *Econometrica*, 27, 575-595.
- Nelson, C.R., and R. Startz (1990): "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator," *Econometrica*, 58, 967-976.
- Pence, K. (2003): "Foreclosing on Opportunity: State Laws and Mortgage Credit," *mimeo*, Board of Governors of the Federal Reserve System.
- Phillips, G.D.A., and C. Hale (1977): "The Bias of Instrumental Variable Estimators of Simultaneous Equation Systems," *International Economic Review*, 18, 219-228.
- Richardson, D.H. (1968): "The Exact Distribution of a Structural Coefficient Estimator," *Journal of the American Statistical Association*, 63, 1214-11226.

- Rothenberg, T.J. (1983): "Asymptotic Properties of some Estimators in Structural Models," in S. Karlin, T. Amemiya and L.A. Goodman (eds.), *Studies in Econometrics, Time Series and Multivariate Statistics*, New York: Academic Press.
- Rothenberg, T.J. (1984): "Approximating the Distribution of Econometric Estimators and Test Statistics," in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Vol. 2, Amsterdam: North Holland.
- Sawa, T. (1969): "The Exact Sampling Distribution of Ordinary Least Squares and Two-Stage Least Squares Estimators," *Journal of the American Statistical Association*, 64, 923-937.
- Staiger, D. and J.H. Stock (1997): "Instrumental Variable Regression with Weak Instruments," *Econometrica*, 65, 557-586.
- Stauffer, H. (1998): *Rechtsprechung des Bundesgerichts zum Sozialversicherungsrecht, Bundesgesetz über die obligatorische Arbeitslosenversicherung und Insolvenzschiädigung*, 2nd edition, Zürich: Schulthess Polygraphischer Verlag.
- Stock, J.H., and M. Yogo (2002): "Testing for Weak Instruments in Linear IV Regression," *NBER Technical Working Paper 284*.
- Stock, J.H., and J.H. Wright (2000): "GMM with weak identification," *Econometrica*, 68, 1055-1096.
- Stock, J.H., J.H. Wright and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments", *Journal of Business and Economic Statistics*, 2002, 518-529.
- Theil, H. (1958): *Economic Forecasts and Policy*. Amsterdam: North Holland.
- Wang, J., and E. Zivot (1998): "Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments," *Econometrica*, 66, 1389-1404.
- Wald, A. (1940): "The Fitting of Straight Lines of Both Variables are Subject to Error", *Annals of Mathematical Statistics*, 11, 284-300.
- Zivot, E., R. Startz and C.R. Nelson (1998): "Valid Confidence Intervals and Inference in the Presence of Weak Instruments," *International Economic Review*, 39, 1119-1144.

Appendix A: Weak instruments estimators and Monte Carlo simulation

In the following Monte Carlo simulation, the finite sample properties of various estimators proposed to deal with the problem of weak instruments are compared for the specific situation in this paper: a binary endogenous regressor and an instrumental variable taking only two different values (within each local labour market). The Monte Carlo is designed to capture some features of the possible effects of ALMP on earnings, in particular, heterogeneous incomes and heterogeneous treatment effects.²⁶

As a versatile but simple design for the Monte Carlo we consider a single local labour market that belongs to 3 different administrative regions: r , s and t . This allows us to consider a single treatment effect θ_{rt} between region r and t and three pair-wise effects $\theta_{rs}, \theta_{rt}, \theta_{st}$ in the same scenario. As explained in Section 3.2, θ_{rt} is the effect on all compliers and is identical to a compliers-weighted average of $\theta_{rs}, \theta_{rt}, \theta_{st}$. In the Monte Carlo simulations, different estimators of $\hat{\theta}_{rt}$ and of the aggregated LATE are examined. This serves two purposes. First, to examine the finite sample properties of the aggregated LATE in different scenarios and for different sample sizes.²⁷ Second, by comparing the properties of $\hat{\theta}_{rt}$ with the aggregated LATE (of the pair-wise effects $\hat{\theta}_{rs}, \hat{\theta}_{rt}, \hat{\theta}_{st}$), the benefits from splitting a single local labour market that stretches into more than two administrative regions into multiple pair-wise comparisons can be assessed (as discussed in Section 3.2).²⁸

²⁶ Since the Monte Carlo simulations are computationally expensive (due to the bootstrapping), only a continuous outcome variable is examined here. We expect the results to be qualitatively similar for a binary outcome variable.

²⁷ Since in this particular set-up the three pair-wise effects $\hat{\theta}_{rs}, \hat{\theta}_{rt}, \hat{\theta}_{st}$ are correlated, the estimation of the weights for the aggregated LATE estimator is more difficult as variances *and* covariances need to be estimated. If all local labour markets were defined as mutually exclusive, all off-diagonal elements of the covariance matrix of $\hat{\theta}$ were known to be zero, yielding less noisy estimates of the optimal weights. The Monte Carlo results can therefore be interpreted as demonstrating the behaviour of the aggregated LATE in a rather difficult scenario.

²⁸ In principle, $\hat{\theta}_{rt}$ is a consistent estimator of the treatment effect on the *total complier subpopulation*, because the impact of the region on the treatment probability is monotonic and the regions r and t are the regions with the low-

The population is spread asymmetrically across these three regions, such that 5% of the population live in region r ($Z_i = z_r$), 65% in region s and 30% in region t . The region of residence affects monotonically the probability of being assigned to treatment D_i but does not affect the potential earnings outcomes Y_i^0 and Y_i^1 . The treatment status is given by

$$D_i = 1(-0.2 \cdot 1(Z_i = z_r) + 0.3 \cdot 1(Z_i = z_t) + u_i > 0) \quad (0.9)$$

and the observed outcome is given by

$$Y_i = \alpha_i + (\beta_i + 1)D_i \quad (0.10)$$

where

$$(\alpha_i, \beta_i, u_i) \sim iidN(0, \Omega). \quad (0.11)$$

α_i are the earnings when not participating in ALMP, $\beta_i + 1$ is the treatment effect. Both are heterogeneous over the individuals. Through the choice of the covariance matrix Ω different average treatment effects, treatment effects on the treated and on the non-treated and different local average treatment effects can be generated. The region of residence shifts the probability of being assigned to treatment, but does not affect the potential outcomes. Observations with $u_i \in (0, 0.2]$ would not participate when living in region r , but would participate in treatment when living in region s . Similarly, observations with $u_i \in (-0.3, 0]$ would change their treatment status if transferred from region s to region t . Hence, the *total complier subpopulation* is all observations with $u_i \in (-0.3, 0.2]$, because observations with $u_i \leq -0.3$ would never participate in treatment and observations with $u_i > 0.2$ would always participate regardless of region of residence. The total complier subpopulation is the largest subpopulation for which a treatment effect is identified.

est and highest, respectively, treatment inclination. However, because the observations in region s are completely neglected in the estimator $\hat{\theta}_{r|t}$, in finite samples a weighted average of $\hat{\theta}_{r|s}$, $\hat{\theta}_{r|t}$, $\hat{\theta}_{s|t}$ might yield a more precise es-

Table A.1 Simulation designs

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6
σ_α^2	64	4	0	64	4	0
σ_β^2	9	9	9	9	9	9
σ_u^2	1	1	1	9	9	9
$\rho_{\alpha\beta}$	0.3	0.3	0	0.3	0.3	0
$\rho_{\alpha u}$	0.5	0.5	0	0.5	0.5	0
$\rho_{\beta u}$	0.9	0.9	0.9	0.9	0.9	0.9
<hr/>						
Fraction treated (in %)		53			51	
Treated in region $r/s/t$		42 / 50 / 62			47 / 50 / 54	
Complier $r-t$ (in %)		20			7	
Complier $r-s$ / $s-t$		8 / 12			3 / 4	
ATE / ATET / ATEN		1.00 / 3.00 / -1.27			1.00 / 3.11 / -1.20	
$\theta_{rt} / \theta_{rs} / \theta_{st}$		0.87 / 1.27 / 0.60			0.96 / 1.09 / 0.87	

Note: ATE is the average treatment effect, ATET is the average treatment effect on the treated. ATEN is the average effect on the non-treated. θ_{rs} is the effect for the subpopulation that would not be treated in region r , but treated in region s . θ_{st} is defined analogously. θ_{rt} is the average treatment effect for the total complier subpopulation.

Six different designs of the covariance matrix Ω are examined, which are summarised in Table A.1. In the first three rows the variances of α_i, β_i, u_i are given and in the subsequent three rows the correlations between these variables. All designs are characterised by a strong correlation between β_i and u_i , which implies that the average treatment effect differs substantially from the treatment effect on the treated, the effect on the non-treated and the local average treatment effect (shown in the lower half of the Table).

In the first design, the variance of the earnings levels is $\sigma_\alpha^2=64$, compared to a variance of the treatment effect of $\sigma_\beta^2=9$. This corresponds to our prior belief, that the variance in earnings is much larger in magnitude than the variance in the treatment impacts. Nevertheless, we also examine scenarios where the variance in earnings is smaller than the variance in impacts (Design 2: $\sigma_\alpha^2=4$ versus $\sigma_\beta^2=9$) and where the variance in earnings is zero (Design 3: $\sigma_\alpha^2=0$ versus $\sigma_\beta^2=9$).

imator.

In all three designs, the size of the total complier subpopulation is 20 %, which is composed of 8% switching to treatment when transferred from region r to s , and the remaining 12% switching from region s to t . Designs 4 to 6 are analogous to the previous designs except for an increase in the variance of the error term u_i in the participation equation. This reduces the power of the instrument and the fractions of compliers between regions r and s and between regions s and t are now only 3% and 4%, respectively.

In the simulations, the OLS, 2SLS, Fuller and 3 jack-knife estimators are examined. Angrist, Imbens and Krueger (1999) and Blomquist and Dahlberg (1999) examined a jack-knifed IV estimator (JIVE) that is motivated by the observation that the first-stage fitted values of the 2SLS estimator are correlated with the error term in the main equation.²⁹ To eliminate this correlation, in a leave-one-out approach the first-stage coefficients are estimated from the sample with observation i removed. The fitted value \hat{D}_i of observation i is estimated as

$$\hat{D}_i = Z_i \left(Z_{-i}' Z_{-i} \right)^{-1} Z_{-i}' D_{-i}, \quad (12)$$

where Z_i is the value of the instrument for observation i and Z_{-i}, D_{-i} are the data matrices (including constants) of the full sample without observation i . Let \hat{D}_N denote the matrix of fitted values \hat{D}_i for all observations. With these fitted values the conventional IV estimator is computed as (JIVE1)

$$\hat{\theta}_N^{JIVE} = \left(\hat{D}_N' D_N \right)^{-1} \hat{D}_N' Y_N.$$

A modification of this jack-knife IV estimator computes the fitted values \hat{D}_i by using the leave-one-out only in the numerator (JIVE2):

$$\hat{D}_i = Z_i \left(Z_N' Z_N \right)^{-1} Z_{-i}' D_{-i} \frac{N}{N-1} \quad (13)$$

²⁹ See also Phillips and Hale (1977).

Hahn, Hausman and Kuersteiner (2002) propose a Jackknife-2SLS (JN2SLS) estimator to eliminate the bias of 2SLS.³⁰ The finite sample bias of 2SLS is estimated through the jack-knife and subtracted from the 2SLS estimator:

$$N \cdot [D_N' P_N D_N]^{-1} D_N' P_N Y_N - \frac{N-1}{N} \sum_{i=1}^N [D_{-i}' P_{-i} D_{-i}]^{-1} D_{-i}' P_{-i} Y_{-i}, \quad (14)$$

where Y_{-i}, D_{-i}, P_{-i} are the data and projection matrices when observation i is removed from the sample.

Both the JIVE and the Jackknife-2SLS estimators possess moments only in the case of overidentification.

The simulation results are given in Tables A.2 to A.7 for the six different designs and for sample sizes 400 and 1600. In the first eight columns, the results for different estimators $\hat{\theta}_{rt}$ are given, where only the observations in regions r and t are used. With respect to mean squared error, the Fuller(4) estimator is always the best, followed by the Fuller(1) estimator, which is slightly worse than OLS when the instrument is very weak and the sample size is small (designs 4 to 6, with 400 observations). With respect to mean absolute error, median squared error and median absolute error, Fuller(1) performs somewhat better than Fuller(4). OLS is always worse than Fuller(1) and Fuller(4). As the estimators 2SLS, JIVE1, JIVE2, JN2SLS have no finite-sample moments, their MSE does not exist. But they also perform clearly worse than Fuller(1) and Fuller(4) with respect to the median error measures, particularly the JIVE estimators.

In the right half of the Tables A.2 to A.7, the results for the aggregated LATE are shown. The complier treatment effects are estimated pair-wise for the three regions by the Fuller estimator and the three estimates $\hat{\theta}_{rs}, \hat{\theta}_{rt}$ and $\hat{\theta}_{st}$ are weighted by different schemes. Six different weighting schemes (with and without bias

³⁰ Donald and Newey (2001) also considered a bias adjusted version of 2SLS where $k = N/(N - L + 2)$ is chosen. In the particular setting considered in this paper, the B2SLS estimator is very similar to the Fuller ($\alpha = 1$) estimator and not further considered here. With one endogenous variable and one instrument ($L=1$), the system is exactly identified. Hence, $k_{LML} = 1$ and the Fuller ($\alpha = 1$) estimator is $[D_N' W D_N]^{-1} D_N' W Y_N$ with $W = (N-2)P_N + I_N$, whereas for B2SLS $W = NP_N + I_N$.

correction) are examined. As most of these weighting schemes depend on estimates of the bias and the covariances of $\hat{\theta} = (\hat{\theta}_{r|s}, \hat{\theta}_{r|t}, \hat{\theta}_{s|t})'$, for each Monte Carlo sample the bias \hat{b} and the covariance matrix $\hat{\Sigma}$ are estimated through 1'000 bootstrap draws. The results for the six different weighting schemes are given separately for the Fuller(1) and for the Fuller(4) estimator.

In the first column, $\hat{\theta}$ is weighted by the estimated complier fraction, and in the second column, the bias corrected estimate $\hat{\theta} - \hat{b}$ is weighted by the complier fraction. The complier fraction for $\hat{\theta}_{r|s}$ is estimated as the difference between the treatment probability in region r and in region s :

$$\frac{\sum D_i \cdot 1(Z_i = z_s)}{\sum 1(Z_i = z_s)} - \frac{\sum D_i \cdot 1(Z_i = z_r)}{\sum 1(Z_i = z_r)},$$

and analogously for $\hat{\theta}_{r|t}$ and $\hat{\theta}_{s|t}$.³¹ In columns 3 and 4, the estimate

$\hat{\theta}$ and the bias corrected estimate $\hat{\theta} - \hat{b}$, respectively, are weighted by \hat{w}_{Var}^{opt} , and in columns 5 and 6 they

are weighted by \hat{w}_{MSE}^{opt} . These optimal weights are estimated as $\hat{w}_{Var}^{opt} = \hat{\Sigma}^{-1} \iota / (\iota' \hat{\Sigma}^{-1} \iota)^{-1}$ and

$$\hat{w}_{MSE}^{opt} = (\hat{\Sigma} + \hat{b}\hat{b}')^{-1} \iota / (\iota' (\hat{\Sigma} + \hat{b}\hat{b}')^{-1} \iota)^{-1}.$$

Tables A.2 to A.7 show that, generally, weighting by the complier fractions leads to a very large MSE. The compliers-weighted average seems to be highly susceptible to the occurrence of zero or near-zero denominators. Turning to the Fuller(4) estimator with \hat{w}_{Var}^{opt} or \hat{w}_{MSE}^{opt} weighting, it can be seen that the inclusion of the region s did not lead to more precise estimates. Compared to using only the estimate $\hat{\theta}_{r|t}$, weighting by \hat{w}_{Var}^{opt} or \hat{w}_{MSE}^{opt} often leads to a substantial deterioration in MSE. This is different for the Fuller(1) estimator, where the results are more favourable to weighting. For all scenarios and sample sizes, weighting by \hat{w}_{Var}^{opt} or \hat{w}_{MSE}^{opt} leads to a lower MSE compared to using only the estimate $\hat{\theta}_{r|t}$. Apart from the designs where earnings do not vary across persons ($\sigma_\alpha^2 = 0$, designs 3 and 6), the aggregated LATE with Fuller(1) has low-

³¹ If $\hat{\theta}_{r|s}, \hat{\theta}_{r|t}, \hat{\theta}_{s|t}$ were estimated by the Wald (=2SLS) estimator, the compliers weighted average would be identical to weighting by $(0, 1, 0)$, because the other terms cancel.

est MSE in all designs and usually performs best also with respect to the other error measures. The reductions in MSE can be up to 30% (in design 1 with sample size 1600). Generally, the different weighting schemes lead to rather similar results, where the weights \hat{w}_{var}^{opt} (without bias correction) yield most often the lowest MSE.

The reason for the different behaviour of the aggregated LATE estimators with Fuller(1) versus Fuller(4) seems to be that Fuller(1) tends to have a higher variance but a lower bias than Fuller(4). The higher variance of the Fuller(1) estimator can be ameliorated through including the observations of region s and weighting the three estimates $\hat{\theta}_{r|s}$, $\hat{\theta}_{r|t}$ and $\hat{\theta}_{s|t}$. This, however, does not seem to work to reduce the bias of the Fuller(4) estimator. Taken together, although no strong conclusions can be drawn from this limited Monte Carlo study, the Fuller(1) estimator with variance-weighting seems to be most suited to our application.

Table A.2: Scenario 1: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (64, 9, 1)$

		$\hat{\theta}_{r t}$						Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)													
								Fuller (1)						Fuller (4)							
		OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	$\hat{w}_{complier}$		\hat{w}_{Var}^{opt}		\hat{w}_{MSE}^{opt}		$\hat{w}_{complier}$		\hat{w}_{Var}^{opt}		\hat{w}_{MSE}^{opt}		
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$
Sample size 400																					
Median	9.09	1.78	6.86	6.22	6.35	4.45	6.67	3.72	3.10	5.73	5.30	5.82	5.48	6.28	5.82	7.76	7.50	7.79	7.58		
Bias	8.23	.	.	4.94	.	3.56	5.63	2.73	1.58	4.67	4.07	4.76	4.24	5.54	5.07	6.72	6.45	6.76	6.51		
Stdc	1.32	6.73	3.68	21.91	26.77	3.78	4.54	3.72	4.33	18.27	21.71	1.99	2.25	1.97	2.20		
MeanSE	69.40	56.59	45.25	487.2	718.8	36.05	37.16	36.45	36.70	364.3	497.0	49.13	46.65	49.61	47.18		
MeanAE	8.26	6.02	5.87	8.38	9.96	5.17	5.13	5.24	5.15	7.24	7.32	6.73	6.47	6.77	6.53		
MedSE	67.77	55.44	.	.	47.70	26.93	33.97	24.02	29.91	24.91	22.55	25.68	23.84	31.10	27.18	47.58	44.04	47.98	45.08		
MedAE	8.23	7.44	11.78	12.00	6.90	5.19	5.82	4.90	5.47	4.99	4.75	5.07	4.88	5.58	5.21	6.90	6.64	6.93	6.71		
Sample size 1600																					
Median	9.08	0.88	-0.59	-0.60	2.02	1.71	3.34	1.91	1.96	2.63	2.67	2.66	2.69	3.54	3.30	5.25	5.02	5.23	5.02		
Bias	8.22	-1.18	-4.61	-7.20	11.79	0.41	2.41	0.48	0.31	1.65	1.64	1.63	1.64	2.34	2.08	4.07	3.86	4.03	3.84		
Stdc	0.65	12.00	468.0	449.4	970.5	5.10	3.50	7.94	8.22	3.15	3.31	3.15	3.28	6.26	6.38	2.39	2.57	2.41	2.57		
MeanSE	68.19	145.8	.	.	.	26.24	18.06	63.16	67.56	12.62	13.64	12.54	13.40	44.57	44.97	22.23	21.52	22.03	21.35		
MeanAE	8.22	5.22	24.09	21.49	31.29	4.06	3.48	3.89	4.19	2.89	3.01	2.87	2.99	3.59	3.61	4.21	4.08	4.19	4.07		
MedSE	67.52	13.16	22.18	21.75	10.26	10.84	9.04	9.11	9.87	6.73	7.18	6.51	7.18	8.58	8.14	19.30	17.34	19.05	17.39		
MedAE	8.27	3.68	4.71	4.66	3.20	3.29	3.01	3.02	3.14	2.59	2.68	2.55	2.68	2.93	2.85	4.39	4.16	4.37	4.17		

Note: First eight columns give results for the complier treatment effect with respect to region r versus region t for OLS, 2SLS, JIVE1, JIVE2, JN2SLS, Fuller(1) and Fuller(4). Observations in region s are not used. 10'000 replications. Values larger 1'000 are not displayed. The subsequent columns provide the results for the aggregated LATE, where $(\hat{\theta}_{r|s}, \hat{\theta}_{r|t}, \hat{\theta}_{s|t})$ are estimated either by Fuller(1) or Fuller(4) and weighted by different weighting schemes. $\hat{w}_{complier}$ refers to weighting by the estimated complier fraction, estimated as $\hat{E}[D|Z=z_s] - \hat{E}[D|Z=z_r]$ for $\hat{\theta}_{r|s}$, and analogously for $\hat{\theta}_{r|t}$ and $\hat{\theta}_{s|t}$. \hat{w}_{Var}^{opt} refers to the optimal weights with respect to the variance of the aggregated LATE, and \hat{w}_{MSE}^{opt} refers to its MSE. $\hat{\theta}$ means that the weighted average of the estimates $(\hat{\theta}_{r|s}, \hat{\theta}_{r|t}, \hat{\theta}_{s|t})$ is taken, whereas $\hat{\theta} - \hat{b}$ refers to the average of the bias corrected estimates. The bias and covariance matrix of $\hat{\theta}$ are estimated (within each Monte Carlo replication) from 1'000 bootstrap draws. Results based on about 1'500 Monte Carlo replications for sample size 400, and 500 replications for sample size 1'600.

Table A.3: Scenario 2: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (4, 9, 1)$

	$\hat{\theta}_{r t}$							Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)														
	OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	Fuller (1)						Fuller (4)								
								$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}		$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}				
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	
Sample size 400																						
Median	4.33	1.32	3.29	2.93	3.27	2.37	3.29	2.31	1.99	3.18	2.98	3.26	3.08	3.23	3.03	4.03	3.92	4.06	3.97			
Bias	3.46	.	.	3.25	.	1.44	2.39	1.63	1.21	2.22	1.98	2.27	2.06	2.46	2.25	3.08	2.97	3.10	3.00			
Stdc	0.42	2.27	1.20	20.27	27.27	1.28	1.53	1.28	1.48	8.38	9.83	0.66	0.75	0.66	0.74			
MeanSE	12.17	7.24	7.17	.	.	6.58	6.27	6.76	6.42	.	.	9.93	9.37	10.05	9.52			
MeanAE	3.46	.	.	24.32	.	2.20	2.42	3.88	4.63	2.30	2.19	2.34	2.24	3.18	3.16	3.08	2.97	3.10	3.00			
MedSE	11.98	6.29	21.82	22.43	7.60	3.51	5.86	3.38	3.99	5.37	4.65	5.77	5.04	5.84	4.89	10.04	9.37	10.20	9.62			
MedAE	3.46	2.51	4.67	4.74	2.76	1.87	2.42	1.84	2.00	2.32	2.16	2.40	2.25	2.42	2.21	3.17	3.06	3.19	3.10			
Sample size 1600																						
Median	4.33	0.90	0.25	0.25	1.40	1.20	1.89	1.32	1.35	1.74	1.75	1.74	1.76	2.05	1.96	2.79	2.67	2.78	2.69			
Bias	3.47	-0.56	-1.55	-5.15	18.01	0.14	1.00	0.31	0.24	0.73	0.71	0.72	0.71	1.07	0.95	1.83	1.74	1.81	1.72			
Stdc	0.21	6.70	.	.	.	1.77	1.13	1.60	1.85	1.13	1.23	1.13	1.21	1.41	1.57	0.88	0.94	0.89	0.94			
MeanSE	12.09	45.16	.	.	.	3.17	2.28	2.65	3.48	1.81	2.00	1.80	1.97	3.13	3.35	4.11	3.89	4.06	3.86			
MeanAE	3.47	1.98	7.45	9.83	23.75	1.39	1.25	1.27	1.41	1.11	1.16	1.10	1.15	1.36	1.35	1.84	1.77	1.83	1.76			
MedSE	12.04	1.62	2.82	2.76	1.25	1.33	1.28	1.29	1.47	1.06	1.22	1.03	1.21	1.55	1.44	3.72	3.27	3.67	3.32			
MedAE	3.47	1.27	1.68	1.66	1.12	1.15	1.13	1.14	1.21	1.03	1.11	1.02	1.10	1.25	1.20	1.93	1.81	1.92	1.82			

Note: See note below Table A.2. Values larger than 100 are not displayed.

Table A.4: Scenario 3: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (0, 9, 1)$

	$\hat{\theta}_{r t}$							Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)																																						
	OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	Fuller (1)						Fuller (4)																																
								$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}		$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}																												
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$																											
Sample size 400																																														
Median	2.75	1.13	2.31	2.16	2.32	1.71	2.21	1.73	1.56	2.30	2.19	2.34	2.25	2.15	2.01	2.74	2.70	2.76	2.71																											
Bias	1.88	.	0.98	5.42	.	0.78	1.31	0.69	0.42	1.37	1.22	1.38	1.26	1.20	1.06	1.85	1.78	1.86	1.79																											
Stdc	0.24	1.36	0.69	7.75	9.57	0.76	0.95	0.76	0.91	3.29	3.77	0.37	0.44	0.37	0.43																											
MeanSE	3.61	2.45	2.21	60.57	91.67	2.44	2.39	2.47	2.40	12.27	15.33	3.55	3.36	3.58	3.40																											
MeanAE	1.88	.	12.57	14.63	.	1.25	1.34	2.05	2.43	1.42	1.37	1.43	1.38	1.71	1.69	1.85	1.78	1.86	1.79																											
MedSE	3.53	2.28	7.39	7.56	2.61	1.23	1.80	1.31	1.57	2.08	1.82	2.18	1.98	1.75	1.50	3.53	3.35	3.57	3.42																											
MedAE	1.88	1.51	2.72	2.75	1.62	1.11	1.34	1.15	1.25	1.44	1.35	1.48	1.41	1.32	1.23	1.88	1.83	1.89	1.85																											
Sample size 1600																																														
Median	2.75	0.89	0.51	0.52	1.19	1.05	1.43	1.09	1.11	1.38	1.36	1.37	1.35	1.50	1.44	2.04	1.95	2.03	1.95																											
Bias	1.89	-0.25	-0.50	-0.32	41.10	0.04	0.52	0.10	0.06	0.45	0.43	0.44	0.42	0.58	0.51	1.10	1.03	1.08	1.02																											
Stdc	0.12	7.53	45.69	55.47	.	1.08	0.68	1.06	1.23	0.74	0.80	0.74	0.79	0.65	0.73	0.54	0.59	0.55	0.59																											
MeanSE	3.57	56.70	.	.	.	1.16	0.75	1.12	1.52	0.74	0.82	0.74	0.81	0.76	0.79	1.49	1.40	1.47	1.39																											
MeanAE	1.89	1.19	5.02	4.50	45.56	0.83	0.71	0.81	0.89	0.72	0.75	0.71	0.74	0.73	0.73	1.12	1.07	1.11	1.06																											
MedSE	3.56	0.57	1.02	0.99	0.43	0.47	0.43	0.48	0.55	0.45	0.49	0.42	0.48	0.47	0.44	1.37	1.19	1.36	1.19																											
MedAE	1.89	0.76	1.01	1.00	0.65	0.69	0.65	0.69	0.74	0.67	0.70	0.65	0.69	0.68	0.67	1.17	1.09	1.16	1.09																											

Note: See note below Table A.2. Values larger than 100 are not displayed.

Table A.5: Scenario 4: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (64, 9, 9)$

	$\hat{\theta}_{r t}$							Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)															
	OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	Fuller (1)						Fuller (4)									
								$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}		$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}					
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$		
Sample size 400																							
Median	9.38	7.45	9.67	9.76	9.29	8.87	9.17	8.25	7.87	8.90	8.65	8.98	8.74	9.08	8.96	9.27	9.22	9.28	9.23				
Bias	8.48	7.75	8.11	7.43	6.91	7.91	7.70	7.96	7.79	8.11	7.99	8.29	8.23	8.30	8.24				
Stdc	1.29	6.37	2.95	32.86	47.42	3.36	4.46	3.16	4.08	17.47	21.74	1.47	1.73	1.43	1.66				
MeanSE	75.04	100.1	74.55	.	.	73.75	79.18	73.33	77.32	370.8	536.7	70.84	70.62	70.92	70.66				
MeanAE	8.48	8.58	8.13	11.89	14.66	7.97	7.94	8.01	7.96	9.51	9.94	8.29	8.23	8.30	8.24				
MedSE	71.19	227.6	104.8	110.9	214.9	64.96	67.59	66.80	81.00	63.38	59.64	64.58	61.12	67.55	66.54	69.31	68.49	69.42	68.66				
MedAE	8.46	15.08	10.23	10.56	14.66	8.05	8.22	8.17	9.00	7.96	7.72	8.04	7.82	8.22	8.16	8.33	8.28	8.33	8.29				
Sample size 1600																							
Median	9.39	3.78	9.42	9.48	8.29	6.95	8.34	6.48	5.66	7.32	6.73	7.50	6.99	7.80	7.45	8.60	8.40	8.66	8.41				
Bias	8.45	.	10.67	11.17	.	5.63	7.05	2.74	0.80	6.33	5.80	6.42	5.96	9.26	9.61	7.64	7.46	7.68	7.51				
Stdc	0.68	.	324.9	824.2	.	6.59	3.18	156.3	211.6	3.68	4.68	3.56	4.39	59.70	73.93	1.44	1.71	1.40	1.64				
MeanSE	71.82	75.18	59.86	.	.	53.55	55.52	53.81	54.76	.	.	60.42	58.49	60.92	59.06				
MeanAE	8.45	.	43.46	59.33	.	7.14	7.12	20.01	25.35	6.56	6.43	6.66	6.51	11.86	12.83	7.64	7.46	7.68	7.51				
MedSE	71.34	107.7	126.7	130.8	91.44	43.08	54.67	39.97	50.21	40.71	35.84	42.94	37.52	49.01	43.69	58.51	55.45	59.43	55.71				
MedAE	8.45	10.39	11.28	11.48	9.56	6.58	7.39	6.32	7.09	6.38	5.99	6.55	6.13	7.00	6.61	7.65	7.45	7.71	7.46				

Note: See note below Table A.2. Values larger than 1'000 are not displayed.

Table A.6: Scenario 5: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (4, 9, 9)$

	$\hat{\theta}_{r t}$							Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)												
	OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	Fuller (1)						Fuller (4)						
								$\hat{w}_{complier}$		\hat{w}_{Var}^{opt}		\hat{w}_{MSE}^{opt}		$\hat{w}_{complier}$		\hat{w}_{Var}^{opt}		\hat{w}_{MSE}^{opt}		
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	
Sample size 400																				
Median	4.60	3.81	4.75	4.82	4.63	4.38	4.51	4.13	4.00	4.46	4.41	4.50	4.43	4.48	4.42	4.63	4.62	4.62	4.62	
Bias	3.66	.	4.55	8.95	.	3.35	3.52	2.99	2.74	3.48	3.39	3.50	3.42	3.36	3.26	3.66	3.64	3.67	3.64	
StdC	0.41	2.07	0.95	16.41	23.32	1.08	1.42	1.02	1.31	6.30	7.75	0.48	0.56	0.47	0.54	
MeanSE	13.58	15.52	13.30	.	.	13.25	13.48	13.28	13.39	50.86	70.69	13.62	13.52	13.65	13.55	
MeanAE	3.66	.	13.81	17.03	.	3.47	3.52	4.99	5.89	3.49	3.42	3.51	3.45	4.00	4.10	3.66	3.64	3.67	3.64	
MedSE	13.36	26.11	17.56	18.41	30.63	11.77	12.64	11.43	11.97	12.35	11.95	12.58	12.13	12.65	12.29	13.54	13.43	13.50	13.49	
MedAE	3.66	5.11	4.19	4.29	5.54	3.43	3.56	3.38	3.46	3.51	3.46	3.55	3.48	3.56	3.51	3.68	3.67	3.67	3.67	
Sample size 1600																				
Median	4.61	2.258	4.64	4.64	4.17	3.49	4.14	3.33	3.07	3.86	3.65	3.93	3.78	3.96	3.80	4.38	4.31	4.40	4.33	
Bias	3.67	.	23.43	7.22	.	2.48	3.08	2.27	1.84	2.83	2.62	2.88	2.70	2.97	2.81	3.40	3.32	3.42	3.35	
StdC	0.20	2.25	1.07	10.73	13.52	1.23	1.59	1.18	1.48	6.10	7.03	0.50	0.60	0.48	0.57	
MeanSE	13.48	11.19	10.62	.	.	9.54	9.41	9.67	9.46	45.96	57.28	11.79	11.40	11.90	11.52	
MeanAE	3.66	.	40.14	31.9	.	2.79	3.08	4.40	5.23	2.86	2.74	2.91	2.79	3.69	3.75	3.40	3.32	3.42	3.35	
MedSE	13.42	13.08	21.07	21.94	15.25	6.88	10.18	7.33	7.41	8.46	7.47	8.91	8.03	9.32	8.54	11.79	11.34	11.90	11.46	
MedAE	3.66	3.62	4.59	4.68	3.91	2.62	3.19	2.71	2.72	2.91	2.73	2.99	2.83	3.05	2.92	3.43	3.37	3.45	3.39	

Note: See note below Table A.2. Values larger than 100 are not displayed.

Table A.7: Scenario 6: $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_u^2) = (0, 9, 9)$

	$\hat{\theta}_{r t}$							Aggregated LATE (weighted average of $\hat{\theta}_{r s}, \hat{\theta}_{r t}, \hat{\theta}_{s t}$)														
	OLS	2SLS	JIVE1	JIVE2	JN2SLS	F(1)	F(4)	Fuller (1)						Fuller (4)								
								$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}		$\hat{W}_{complier}$		\hat{W}_{Var}^{opt}		\hat{W}_{MSE}^{opt}				
								$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	$\hat{\theta}$	$\hat{\theta} - \hat{b}$	
Sample size 400																						
Median	3.01	2.62	3.10	3.13	3.02	2.90	2.96	2.83	2.74	2.96	2.92	2.98	2.94	2.96	2.92	3.06	3.04	3.06	3.05			
Bias	2.07	.	2.49	2.52	.	1.93	2.01	1.88	1.80	2.01	1.96	2.03	1.98	1.93	1.88	2.11	2.10	2.11	2.10			
StdC	0.25	1.25	0.56	6.81	9.57	0.65	0.88	0.61	0.80	2.75	3.35	0.28	0.34	0.27	0.32			
MeanSE	4.73	5.29	4.34	49.90	94.79	4.47	4.61	4.47	4.56	11.30	14.77	4.54	4.51	4.55	4.51			
MeanAE	2.07	.	8.16	11.41	.	2.01	2.01	2.80	3.30	2.02	1.98	2.03	2.00	2.28	2.33	2.11	2.10	2.11	2.10			
MedSE	4.25	9.64	5.71	6.05	11.12	3.81	4.06	3.89	4.14	4.05	3.88	4.12	3.97	4.13	4.00	4.44	4.38	4.46	4.40			
MedAE	2.06	3.11	2.39	2.46	3.33	1.95	2.01	1.97	2.03	2.01	1.97	2.03	1.99	2.03	2.00	2.11	2.09	2.11	2.10			
Sample size 1600																						
Median	3.02	1.71	3.04	3.04	2.77	2.40	2.75	2.31	2.17	2.65	2.53	2.69	2.60	2.68	2.63	2.95	2.93	2.97	2.93			
Bias	2.07	.	0.85	3.22	.	1.40	1.74	1.22	0.77	1.69	1.58	1.72	1.62	1.77	1.65	1.99	1.95	2.00	1.96			
StdC	0.12	1.35	0.63	29.34	39.45	0.76	0.98	0.73	0.91	13.24	16.14	0.33	0.40	0.32	0.38			
MeanSE	4.30	3.82	3.43	.	.	3.43	3.44	3.48	3.47	.	.	4.07	3.97	4.10	4.00			
MeanAE	2.07	.	11.06	11.50	.	1.62	1.74	4.74	5.98	1.72	1.66	1.75	1.69	3.10	3.36	1.99	1.95	2.00	1.97			
MedSE	4.29	4.33	6.77	7.13	4.76	2.28	3.22	2.61	3.14	2.89	2.56	3.04	2.74	3.29	3.05	4.02	3.90	4.06	3.94			
MedAE	2.07	2.08	2.60	2.67	2.18	1.51	1.80	1.62	1.77	1.70	1.60	1.74	1.66	1.82	1.75	2.00	1.97	2.01	1.98			

Note: See note below Table A.2. Values larger than 100 are not displayed.

Appendix B: Data and Institutions

B.1 AVAM/ASAL/AHV Database

The empirical analysis is based on data from two administrative sources, matched by a unique person identifier. Data were obtained for all of the approximately 180'000 persons registered unemployed on December 31, 1997. The first data source is the information system for placement and labour market statistics (AVAM) and the unemployment offices payment system (ASAL). These data provide detailed information about the unemployment history, ALMP participation and personal characteristics for the period 1996 to 1999.³² These data were merged with the social security records (AHV) for a random subsample of about 80'000 individuals. These social security data contain monthly information on the individual labour market histories and earnings (including self-employment) on a monthly basis for the period 1988 to 1999. These social security data allow us to examine the effects of active labour market programmes on employment as well as on earnings.

Sample selection

In order to focus on the mobile population, as described in Section 3, various sample selection rules were applied. These are detailed in Table B.1 together with the ensuing loss of observations.

³² E.g. information on place of residence, mother tongue, foreign languages, number of dependent persons, disability, education, qualification, caseworker's assessment of employability, information about last job, looking for part-time or full-time job, beginning and end of entitlement period, benefit payments, sanction days, sickness benefits, participation in labour market programmes, and temporary wage subsidies. AVAM: Arbeitsvermittlung und Arbeitsmarktstatistik, ASAL: Auszahlungssystem der Arbeitslosenkassen, AHV: Alters- und Hinterlassenenversicherung. The data were merged by Jonathan Gast (seco).

Table B.1: Sample selection

	Remaining observations
Original sample	81'399 obs
Delete all students, apprentices, homeworkers	80'775 obs
Delete all foreigners with less than yearly work permit	79'610 obs
Delete age < 25 or age > 55	75'992 obs
Delete if insured monthly earnings < 1000 CHF	70'474 obs
Delete if partly employed	68'190 obs
Delete if disabled or exhausted unemployment benefits (ausgesteuert) → Final sample	66'713 obs

Note: 66'713 observations. ^a 1 CHF ≈ 2/3 Euro.

Table B.2 shows descriptive statistics for the treated and the non-treated. 60% of all unemployed entered active labour market programmes (with the average month of entry being March 1998). Generally, the characteristics of treated and non-treated are very similar except for slightly higher earnings of the treated in the years up to 1997 and worse outcomes in 1998. The aggregate figures for the treated, however, hide substantial variation between the participants in training, employment programmes and temporary wage subsidies, as can be seen in Gerfin and Lechner (2002). 34% of the treated received training, 22% participated in an employment programme, 42% became employed in a temporary wage subsidy, and the remaining 1% participated in other labour market programmes.

B.2 Regional employment offices and descriptive statistics

Table B.3 gives a list of all regional employment offices. The first column contains an identification number. The second column indicates the canton, and the third column gives the name of the regional employment office. In the last column, the number of observations in the selected sample (see Table B.1) is given. The REO that are included in any of the 30 local labour markets are marked in bold.

Table B.2: Descriptive statistics (Means of shares)

Variable name		Treated	Non-treated
Age		38.2	37.3
Female		45	41
Married		59	59
Number of dependent persons		2.5	2.4
Work permit (for foreigners):	yearly permit (B)	16	16
	residence permit (C)	28	29
	Swiss national	56	55
Mother tongue:	German	34	34
	French	20	20
	Italian	11	11
	Other	35	36
Qualification:	occupation specific skills	56	56
	some occupation specific skills	15	15
	no occupations specific skills	29	29
Assessment of potential employability: by caseworker	very good	5	7
	good	16	16
	medium	56	55
	difficult	16	15
	very difficult	2	2
Insured earnings (CHF) ^a		3'980	3'832
Employment history 1988-1997			
Number of previous unemployment spells		2.0	2.2
Average duration of unemployment spells (months)		5.4	5.1
Number of employment spells		2.5	2.7
Average duration of employment spells (months)		50.6	47.2
Has been self-employed some time in 1993-1997		5.0	5.4
Average monthly earnings from employment	1988-1989	2'682	2'557
	1990-1991	3'081	2'988
	1992-1993	3'072	2'985
	1994-1995	3'175	3'092
	1996-1997	2'757	2'620
Treatment status in 1998			
Month of entry		March 1998	
	of which received training	34	-
	of which participated in employment programme	22	-
	of which engaged in a temporary wage subsidy	42	-
	of which entered in other labour market programmes	1	-
Outcome variables			
Employed	average for January-April 1999	45	53
	average for May-August 1999	59	63
	average for September-December 1999	66	64
High earnings (> 90% of insured earnings)	average for January-April 1999	27	35
	average for May-August 1999	35	42
	average for September-December 1999	37	42
Labour earnings (from wage and self-employment) in CHF	average for January-April 1999	1'714	2'091
	average for May-August 1999	2'218	2'455
	average for September-December 1999	2'408	2'470
Observations		40'193	26'520

Note: 66'713 observations. ^a 1 CHF ≈ 2/3 Euro.

Table B.3: Regional employment offices (REO)

No	Can ton	Name of regional office	Obs	No	Can ton	Name of regional office	Obs
1	ZH	Opfikon	733	51		Kriens	485
2		Marthalen	115	52		Sursee	399
3		Winterthur	1221	53		Wolhusen	154
4		Effretikon	320	54	UR	Altdorf	150
5		Uster	928	55	SZ	Goldau	337
6		Wetzikon	719	56		Lachen	529
7		Bülach-Nord	305	57	OW	Hergiswil, OW	110
8		Meilen	474	58	NW	Hergiswil, NW	155
9		Thalwil	947	59	GL	Glarus	129
10		Affoltern a.A.	277	60		Schwanden	135
11		Dietikon	767	61	ZG	Zug	571
12		Regensdorf	393	64	FR	Châtel-St.Denis	107
13		Ausstell.	607	65		Murten, See-B.	197
14		Ausstell.	537	198		FRD2	2
15		Flösserg.	471	66		Tafers, Sense-	201
16		Flösserg.	360	67		Romont, Glâne	143
18		Schwamend ^a	291	68		Fribourg, Sar.	860
19		Hohlstr.	291	200		FRL2	5
20		Bäckerstr.	615	205		FRL3	8
21		Badenerstr	511	212		Städtisches Arbeitsamt Fribourg	55
22		Eggbühlstr	810	69		Bulle,Gruyère	319
23		Lagerstr.	492	201		FRM2	2
216	BE	BEAa	1	70		Estavayer le lac, Bro	228
27		Mittell. Gümlingen	416	71	SO	Trimbach	104
28		Mittell. Zollikofen	369	74		Grenchen	191
29		Mittell. Liebefeld-Köniz	456	75		Solothurn	440
30		BE-Oberl. Unterseen	366	76		Oensingen	181
214		BEAG	16	77		Olten	336
32		Spiez	223	78		Breitenbach	159
33		BE-Oberl. Meiringen	79	79		Biberist	125
35		Emmental, Langnau	94	80		Schönenwerd	131
36		Oberaargau, Wangen	202	81		Zuchwil	131
37		Lyss (Seeland)	329	196	BS	BSA2	3
38		Moutier	124	82		Basel, RAV 1, Hochstr	783
39		Jura La Neuveville	49	83		Basel, Utengasse	760
40		Jura St.Imier	91	208		KIGA Basel, LAM	56
41		SubRav Bern-West Lagerhausweg	813	84		Basel, RAV 2, Hochstr	538
42		SubRav. Bern-Ost Kasterhoferstr.	666	194	BL	BLA1	14
215		BEAZ	2	87		Pratteln	338
43		Jura, Tavannes	123	197		BLD2	1
204		BEA3	2	88		Münchenstein	256
44		Biel	945	199		BLE2	2
210		BEA6	3	89		Binningen	340
45		Burgdorf	303	90		Liestal	287
46		Langenthal	313	91		Laufen	86
47		Oberl. Thun S.	858	92		Gelterkinden	155
48	LU	Luzern	538	93	SH	Schaffhausen	605
49		Emmen	429	94	AR	Herisau	247
50		Emmenbrücke	155	95	AI	Appenzell	39

193	SG	SGA1	1	140	Echallens	163
96		St.Gallen	1022	141	Vevey	654
97		Rorschach	348	142	Prilly	263
98		Oberuzwil	505	143	Aigle	515
99		Sargans	207	144	Morges	442
100		Buchs	168	145	Renens	801
101		Heerbrugg	417	146	Montreux	411
102		Rapperswil	360	147	Oron-la-Ville	107
103		Wattwil	185	148	Orbe	406
104	GR	Chur	289	149	Pully	378
189		Gemeindearbeitsamt Stadt Chur	191	150	Aubonne	158
105		Thuisis	80	151	Yverdon Grand.	617
106		Illanz	70	153	Nyon	576
107		Grono	73	154	Payerne	236
108		Davos-Platz	76	155	Moudon	119
109		Samedan	98	156	VS Monthey 1	367
110	AG	Zofingen	472	157	Monthey 2	242
111		Zurzach	214	158	Martigny 1	336
112		Muri AG	175	159	Martigny 2	389
113		Menziken	221	160	Fully	123
114		Frick	106	161	Sierre 1	372
192		AGA1	3	162	Sierre 2	319
115		Aarau, Kunsth.	551	163	Sion 1	238
116		Baden	572	164	Sion 2	308
117		Wettingen	484	165	Sion 3	588
118		Lenzburg	328	166	St-Séverin	114
119		Brugg	283	167	Brig 1	539
120		Wohlen	473	168	Brig 2	128
121		Rheinfelden	233	169	Visp	89
122	TG	Frauenfeld	537	170	Naters	7
202		Gemeindearbeitsamt Stadt Frauenfeld	29	171	NE Mont., La Chaux de Fonds	845
123		Kreuzlingen	387	172	Litt., Neuchatel	1234
207		TGP4	1	211	NEA6	1
124		Weinfelden	227	213	Städtisches Arbeitsamt Neuchatel	17
125		Amriswil	474	179	GE Ag de Rive	1509
203		TGQ2	3	180	GE, Carouge	231
206		TGQ3	3	181	AG d Minot. I	1304
209		TGQ5	11	182	AG d Minot. II	613
126	TI	Bellinzona	697	183	AG des Saules	699
127		Locarno	1495	184	AG des Gares	1344
128		Lugano	1665	191	GEA1	2
129		Chiasso	394	195	JU JUA1	1
130		Biasca	343	186	Delémont	266
131	VD	Lausanne Chaud. 9	2010	187	Porrentruy	237
132		Lausanne Chaud. 7	140	188	Saignelégier	46

Note: All regional employment offices with at least one observation in the selected sample. Marked in **bold** are all REO included in any of the 30 local labour markets. ^a Oberwiesenstrasse 66

Appendix C: Local labour markets

Table C.1: Commuting times by car between regional employment offices (in minutes)

	122 Frauenfeld	123 Kreuzlingen	124 Weinfelden	125 Amriswil	55 Goldau	56 Lachen	94 Herisau	95 Appenzell	104 Chur	01 Opfikon	03 Winterthur	05 Uster	06 Wetzikon	08 Meilen
96 St.Gallen				20			16	24						
97 Rorschach				21										
98 Oberuzwil	27		35	36			22				34			
99 Sargans									22					
100 Buchs														
101 Heerbrugg														
102 Rapperswil						11				34	43	24	19	27
103 Wattwil													41	

	122 Frauenfeld	123 Kreuzlingen	124 Weinfelden	125 Amriswil	55 Goldau	56 Lachen	93 Schaffhausen	61 Zug	110 Zofingen	111 Zurzach	112 Muri	113 Menziken	114 Frick	115 Aarau	116 Baden	117 Wettingen	118 Lenzburg	119 Brugg	120 Wohlen	121 Rheinfelden
01 Opfikon	30						34			34	38			37	21	21	30	41	33	
02 Marthalen	25						12													
03 Winterthur	20						22													
04 Effretikon	26						29													
05 Uster	35																			
06 Wetzikon																				
07 Bülach-Nord										29					35	35				
08 Meilen																				
09 Thalwil																				
10 Affoltern																			31	30
11 Dietikon										38	30									
12 Regensdorf															15	15	24	27	24	
13-23: ZH City															19	19	28	39	31	
															27	27		45	37	

	82-84 Basle	87 Pratteln	88 München- stein	89 Binnin- gen	90 Liestal	91 Laufen	92 Gel- terkin- den	186 Delémont	110 Zofingen	111 Zurzach	112 Muri	113 Menziken	114 Frick	115 Aarau	116 Baden	117 Wettin- gen	118 Lenzburg
87 Pratteln	10																
88 Münchenstein	13																
89 Binningen	7																
90 Liestal	20																
91 Laufen	25							20									
92 Gelterkinden	24																
71 Trimbach					28		25		21					20			35
72 Derendingen									27					43			
74 Grenchen																	
75 Solothurn									31								
76 Oensingen							32		17				34				
77 Olten							32		16				36	15			30
78 Breitenbach	26		22			5		26									
79 Biberist									31								
80 Schönenwerd									20				5				24
81 Zuchwil									26								

	68 Fri- bourg	64 Châtel- St.Deni- s	65 Murten	66 Tafers	67 Romont	69 Bulle	70 Esta- vayer le-lac	71 Trim- bach	72 Der- endin- gen	74 Grench- en	75 Solo- thurn	76 Oensin- gen	77 Olten	78 Breiten- bach	79 Biberist	80 Schöne- nwerd	81 Zuchwil	110 Zofin- gen	171 La- Chaux de Fonds	172 Neucha- tel
27 Gümligen	29		26	29					26	42	32	34			27		34			
28 Zollikofen	30		27	30					22	29	29	30			24		30			
29 Liebefeld- Köniz	24		24	25					30	46					31					
30 Unterseen																				
47 Thun																				
32 Spiez																				
33 Meiringen																				
44 Biel			39						25											30
35 Langnau																				
36 Wangen								30	9	28	13	8	24		13	32	14			
37 Lyss			29						30	27	36				36					
38 Moutier																				
39 La Neuveville																				
40 St.Imier																			17	15
41/42 Bern	23		20	24					26	42	32				27		34			
43 Tavannes																				
45 Burgdorf											28				22		29			
46 Langenthal								36	25		29	19	30		29	39	31	26		

Table C.1 shows the commuting distances between the regional employment offices as calculated using <http://route.web.de/Europa>. With these commuting distances and further considerations on commuting infrastructure, 30 local labour markets were identified. These are displayed in Table C.2, together with the number of observations in the selected sample. For example, the first row indicates that the areas corresponding to the REOs of Grenchen, Biel and Lyss are considered as a local labour market, where Grenchen belongs to the canton Solothurn (SO) and Biel and Lyss belong to the canton Bern (BE). This row further shows that 186 observations belong to the REO Grenchen, while the REOs Biel and Lyss together have 1'109 observations.

Table C.2: Local labour markets divided by administrative border

#	Cantons	Regional employment offices		Number of observations	
				n ₁	n ₂
1	SO-BE	Grenchen	Biel, Lyss	186 ^a	1'109 ^a
2	SO-BE	Solothurn, Oensingen, Biberist, Zuchwil	Wangen, Langenthal, Burgdorf	877	818
3	SO-BL	Breitenbach	Laufen	159	86
4	SO-AG	Trimbach, Olten, Schönenwerd	Zofingen, Aarau	571	1'023
5	NE-BE	La Chaux de Fonds	St.Imier	845	91
6	NE-BE	Neuchatel	La Neuveville	1'234	49
7	BE-AG	Langenthal	Zofingen	313	472
8	BE-FR	Gümligen, Zollikofen, Köniz, Bern (2x)	Murten, Tafers, Fribourg	2'660 ^a	763 ^a
9	FR-VD	ChatelSt.Denis	Oron la Ville	107	107
10	FR-VD	Romont, Estavayer	Payerne, Moudon	371	355
11	VD-GE	Nyon	Genf (6x)	576	5'700
12	VD-VS	Vevey, Aigle, Montreux	Monthey (2x)	1'580	609
13	BL-BS	Pratteln, Münchenstein, Binningen	Basel (3x)	934	2'081
14	LU-AG	Sursee	Zofingen	399	472
15	LU-NWOW	Luzern, Emmen, Emmenbrücke, Kriens	Hergiswil (2x)	1'607	265
16	LU-ZG	Luzern, Emmen, Emmenbrücke, Kriens	Zug	1'607	571
17	SZ-UR	Goldau	Altdorf	337	150
18	SZ-ZG	Goldau	Zug	337	571
19	AG-ZH	Baden, Wettingen, Wohlen	Opfikon, Effretikon, Uster, Wetzikon, Bülach, Dietikon, Regensdorf	1'529	4'165
20	ZH-SH	Marthalen	Schaffhausen	115	605
21	ZH-TG	Winterthur	Frauenfeld	1'221	537
22	ZH-SG	Meilen, Thalwil	Rapperswil	1'421	360
23	ZH-SZ	Meilen, Thalwil	Lachen	1'421	529
24	TG-SH	Frauenfeld	Schaffhausen	537	605
25	TG-SG	Amriswil	Rohrschach, Oberuzwil	474	853
26	SG-AR	St.Gallen, Rohrschach	Herisau	1'370	247
27	AR-AI	Herisau	Appenzell	247	39
28	SG-SZ	Rapperswil	Lachen	360	529
29	SG-GR	Sargans	Chur (2x)	207	480
30	GR-TI	Grono	Bellinzona	73	697

Note: ^a Number of observations after deleting individuals with French mother-tongue, because a French-German bilingual region is bordering a German speaking region.

Table C.3 shows various summary statistics for these local labour markets. The first two columns indicate the number of the local labour market and the cantonal border that passes through this labour market. Columns three to six are analogous to Table 4.2 and give the percentages treated, the estimated fraction of compliers and the differential in the quota per unemployed (as given by Table 2.1).

Columns 7 to 14 show the ALMP-structure in the REOs of the local labour market, where columns 7 to 10 refer to the one side of the border and columns 11 to 14 to the other side. The ALMP-structure is presented by the percentages of treated participating in training, employment programmes, temporary wage subsidies and other programmes, respectively. These columns indicate a non-negligible variation in the types of programmes to which caseworkers assign their clients. For example, the first row indicates that 13% of the participants in the REO Grenchen receive training, while this figure is 29% in the REOs Biel and Lyss. The corresponding fractions for employment programmes are 46% in Grenchen and 31% in Biel and Lyss. Hence, not only are unemployed living in Grenchen (compared to Biel and Lyss) more likely to be treated (76% vs. 67%), they are also more likely to be assigned to employment programmes if treated and less likely to receive training. Thus, the treatment itself is different in Grenchen than in Biel and Lyss, which may violate assumption 2 of Section 3.1. The criterion 4 of Section 4.3 therefore requires that the ALMP should be similar on both sides of the border. With the fractions of the ALMP allocation adding up to 100%, a convenient approach to implement Criterion 4 is to limit the *maximum divergence* between the corresponding fractions on the two sides of the border. For example, the maximum divergence in the first row is 16 (29% training - 13% training). For the second labour market, the maximum divergence is 4 (39%-35% for employment programmes). Since the choice of a threshold for the maximum divergence is somewhat arbitrary, two variants are considered: In the first variant (Selection A) a maximum divergence of at most 7 is permitted. This condition is satisfied by only 8 local labour markets, which are marked in the second last column of Table C.3.³³ For the second variant (Selection B), the threshold is doubled to include

³³ The threshold 7 was chosen, as this is exactly the maximum divergence for four of the 30 labour markets. Hence, a smaller threshold would have been satisfied only by four or fewer labour markets.

a larger selection of local labour markets in the analysis.³⁴ Selection B contains 18 local labour markets, marked in the last column of Table C.3. For the estimation of the aggregated LATE only the local labour markets classified as A or B will be used henceforth, since for the other labour markets the conditions for identifying the effect of ALMP as a whole do not seem to be satisfied.³⁵

Table C.3: Treatment intensity and ALMP composition in the local labour markets

Region	% Treated	Com- plier ^a	Change in Instrument Table 2.1 ^b	Allocation to programme types among the treated (in %)								S ^c			
				Train- ing	Empl. prog	Temp subsidy	other	Train- ing	Empl. prg	Temp subsidy	other	A	B		
1	SO-BE	76	67	8.9	-3.9	13	46	39	1	29	31	40	0		
2	SO-BE	68	63	5.6	-3.9	20	35	45	1	18	39	43	0	x	x
3	SO-BL	52	63	-11.2	-4.8	29	1	66	4	24	24	52	0		
4	SO-AG	64	68	-3.4	-5.1	19	29	51	1	34	30	35	0		
5	NE-BE	55	65	-10.0	-3.2	21	42	36	1	27	27	46	0		
6	NE-BE	59	67	-8.4	-3.2	23	39	37	1	45	21	30	3		
7	BE-AG	64	68	-4.9	-1.2	17	37	46	0	24	39	36	1		x
8	BE-FR	65	67	-1.7	-0.2	28	33	39	0	28	38	33	1	x	x
9	FR-VD	74	59	15.0	3.0	30	20	48	1	27	16	56	2	x	x
10	FR-VD	64	60	3.3	3.0	31	30	37	2	21	38	40	1		x
11	VD-GE	57	50	6.5	0.8	41	12	46	1	40	10	48	2	x	x
12	VD-VS	59	66	-6.7	-0.7	33	20	46	1	23	33	43	1		x
13	BL-BS	67	53	14.2	2.1	34	22	44	0	25	22	52	1		x
14	LU-AG	62	68	-6.3	-1.2	46	25	27	1	24	39	36	1		
15	LU-NWOW	64	62	2.4	-10.4	46	21	32	1	39	21	40	1		x
16	LU-ZG	64	64	-0.2	-1.5	46	21	32	1	38	23	38	1		x
17	SZ-UR	69	61	8.8	-1.3	41	38	21	0	34	35	31	0		x
18	SZ-ZG	69	64	5.0	4.7	41	38	21	0	38	23	38	1		
19	AG-ZH	64	58	6.6	3.5	49	15	35	1	38	14	47	1		x
20	ZH-SH	53	63	-9.6	1.2	31	15	54	0	46	18	36	0		
21	ZH-TG	59	69	-9.9	-1.8	38	19	43	1	34	26	40	0	x	x
22	ZH-SG	56	60	-3.8	-4.5	51	12	35	1	45	13	42	0	x	x
23	ZH-SZ	56	72	-15.2	-8.5	51	12	35	1	49	20	31	0	x	x
24	TG-SH	69	63	6.3	3.1	34	26	40	0	46	18	36	0		x
25	TG-SG	64	66	-1.5	-2.7	44	19	37	0	44	15	39	1	x	x
26	SG-AR	61	64	-2.9	-5.1	47	12	40	1	8	38	54	0		
27	AR-AI	64	64	-0.5	-2.6	8	38	54	0	56	0	44	0		
28	SG-SZ	60	72	-11.4	-3.9	45	13	42	0	49	20	31	0		x
29	SG-GR	65	59	5.6	2.3	48	12	39	1	27	20	50	2		
30	GR-TI	52	57	-4.6	3.4	42	21	34	3	25	27	47	2		

Note: ^a The estimate of the fraction of compliers is the difference between the previous two columns.

^b Difference in the instrument quota per unemployed (Table 2.1, column 5) between the two cantons.

^c Selection A contains only labour markets with maximum divergence between the programme type compositions ≤ 7 . For selection B the maximum divergence is ≤ 13 .

³⁴ The threshold was set to 13, since there was no local labour market with maximum divergence of 14.

³⁵ If we were to impose more structural assumptions, these labour markets could also be useful for disentangling the effects of training, employment programmes and temporary wage subsidies.

Table C.4 shows the municipality sizes for the local labour markets, which are used to define urban regions in Section 5.3. In the columns on the left, the *average* municipality size is given for each local labour market, while the columns on the right show the *maximum* municipality size. The minimum is not shown, as it is almost always 500 residents. These sizes are calculated from the municipalities belonging to the REO in this labour market, separately for the two sides of the border. The average size is computed as the size of the municipality weighted by the number of observations (in the dataset) residing in this municipality. Hence, it is self-weighted by the unemployed population. Since the size of the municipality was coded in intervals, the interval-midpoints are used as an approximation to the true municipality size (see note below Table C.4).

Table C.4: Size of municipalities in the local labour markets

		Average municipality size (1'000 residents)		Maximum municipality size (1'000 residents)	
		In the one canton	In the other canton	In the one canton	In the other canton
1	SO-BE	12.8	37.0	15.0	75.0
2	SO-BE	6.3	6.5	15.0	15.0
3	SO-BL	3.0	2.0	7.5	3.5
4	SO-AG	7.3	7.1	15.0	15.0
5	NE-BE	22.9	4.3	40.0	7.5
6	NE-BE	20.2	2.3	40.0	3.5
7	BE-AG	7.6	5.8	15.0	7.5
8	BE-FR	81.2	16.9	150.0	40.0
9	FR-VD	1.9	0.6	3.5	1.5
10	FR-VD	1.5	4.2	3.5	7.5
11	VD-GE	6.5	85.3	15.0	150.0
12	VD-VS	9.2	8.8	15.0	15.0
13	BL-BS	12.9	143.9	15.0	150.0
14	LU-AG	3.0	5.8	7.5	7.5
15	LU-NWOW	35.4	4.0	75.0	7.5
16	LU-ZG	35.4	12.6	75.0	25.0
17	SZ-UR	7.9	4.1	15.0	7.5
18	SZ-ZG	7.9	12.6	15.0	25.0
19	AG-ZH	8.0	12.6	15.0	300.0
20	ZH-SH	1.6	26.3	3.5	40.0
21	ZH-TG	60.8	5.6	300.0	15.0
22	ZH-SG	12.4	7.4	300.0	15.0
23	ZH-SZ	12.4	5.2	300.0	7.5
24	TG-SH	5.6	26.3	15.0	40.0
25	TG-SG	7.6	7.8	15.0	15.0
26	SG-AR	44.9	7.3	75.0	15.0
27	AR-AI	7.3	3.3	15.0	3.5
28	SG-SZ	7.4	5.2	15.0	7.5
29	SG-GR	4.2	20.2	7.5	40.0
30	GR-TI	0.9	8.4	1.5	15.0

Note: Since municipality size was coded in brackets, interval midpoints are used in the calculations: 500 (for bracket <1'000 residents), 1'500 (for bracket 1'000-2'000 residents), 3'500 (for 2'000-5'000 residents), 7'500 (for 5'000-10'000), 15'000 (for 10'000-20'000), 25'000 (for 20'000-30'000), 40'000 (for 30'000-50'000), 75'000 (for 50'000-100'000), 150'000 (for 100'000-200'000), 300'000 (for > 200'000).

Appendix D: Additional estimation results

Table D.1 displays the estimation results for the 30 local labour markets when using the Wald estimator, instead of the Fuller estimator (as in Table 5.1).

Table D.1: Wald estimates of the local average treatment effects for the 30 local labour markets

	Region	Wald estimates for year 1999					
		Employment (in %-points)		High earnings (in %-points)		Labour earnings in CHF	
		May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.
1	SO-BE	3	-14	-27	-40	940	180
2	SO-BE	-166	-114	-77	-50	-5'400	-3'270
3	SO-BL	18	34	0	-9	-430	-200
4	SO-AG	-4	-57	-15	-40	-3'350	-5'960
5	NE-BE	51	69	75	96	2'010	940
6	NE-BE	0	-6	14	24	710	-740
7	BE-AG	-261	-191	-62	-41	-5'050	-3'460
8	BE-FR	-244	-145	12	68	-5'550	-3'070
9	FR-VD	92	44	66	47	3'500	1'960
10	FR-VD	59	119	132	150	7'840	10'640
11	VD-GE	108	163	62	89	8'420	10'140
12	VD-VS	88	87	54	39	1'390	1'090
13	BL-BS	-5	5	21	10	1'660	1'710
14	LU-AG	-152	-114	-129	-119	-6'550	-6'050
15	LU-NWOW	-340	-414	-188	-304	-20'170	-25'590
16	LU-ZG	-2'406	-1'426	-103	678	113'490	131'530
17	SZ-UR	-20	-80	20	-1	-1'520	-4'680
18	SZ-ZG	8	-36	-21	-67	-8'790	-10'680
19	AG-ZH	17	16	57	71	-260	-110
20	ZH-SH	-40	-73	-26	-52	-6'130	-7'300
21	ZH-TG	38	17	2	-11	210	-470
22	ZH-SG	6	-105	4	-74	-7'420	-10'830
23	ZH-SZ	16	9	1	-2	-680	-1'020
24	TG-SH	66	80	15	34	2'650	2'660
25	TG-SG	-8	-104	-8	-154	6'240	-290
26	SG-AR	50	2	53	74	8'210	10'480
27	AR-AI	1'400	206	187	-431	38'700	12'730
28	SG-SZ	20	47	0	22	1'600	2'290
29	SG-GR	-110	-134	-58	-28	-6'700	-4'820
30	GR-TI	-290	-40	-179	-24	-12'800	-4'250

Note: See note below Table 5.1.

Table D.2 gives the estimation results for the aggregated LATE for Selection A (analogously to Table 5.2). The first rows give the estimated weights, which were estimated separately for each outcome variable. The local labour markets 21 and 23 receive about 70% of the weights, while the remaining weights are spread relatively evenly among the other labour markets.³⁶ The estimation results indicate a positive effect of participation in ALMP on employment prospects. The probability of being employed is increased by 14 to 26 percentage points, with only the larger of these two estimates being significantly different from zero. Also the effect on being employed with at least 90% of previous earnings is positive, but insignificant. In Table

³⁶ The covariance matrix $\hat{\Sigma}$ is estimated through 10'000 bootstrap replications, and its inverse is shown in Table D.3.

D.2 the Quantiles $Q_{0.005}$, $Q_{0.025}$ and $Q_{0.05}$ are shown instead of $Q_{0.95}$, $Q_{0.975}$ and $Q_{0.995}$ when the estimated effect is negative.

Table D.2: Aggregated treatment effect for compliers, Selection A

	Employment (in %-points)		Outcome variables, year 1999 High earnings (in %-points)		Labour earnings in CHF	
	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.	May-Aug.	Sept.-Dec.
weights (%)						
BE-FR	2	4	7	6	6	8
FR-VD	8	9	7	9	11	11
VD-GE	9	6	11	10	4	4
ZH-TG	19	21	19	20	24	23
ZH-SG	5	2	5	4	2	1
ZH-SZ	50	52	44	47	46	44
TG-SG	8	6	8	5	7	10
Estimate	26.4	14.2	10.3	6.0	214	-250
Std.error	11.3	10.9	10.8	11.4	629	612
t-value	2.33	1.30	0.95	0.53	0.34	0.41
		Bootstrap of estimate with fixed weights				
$Q_{0.05} / Q_{0.95}$	20.6	21.5	18.2	20.2	1'061	-917
$Q_{0.025} / Q_{0.975}$	<u>24.5</u>	25.0	21.7	24.0	1'255	-1'126
$Q_{0.005} / Q_{0.995}$	32.1	32.5	29.9	31.5	1'668	-1'566
		Bootstrap of estimate with stochastic weights				
$Q_{0.05} / Q_{0.95}$	14.3	16.5	12.6	14.9	767	-849
$Q_{0.025} / Q_{0.975}$	16.2	20.4	17.0	17.7	966	-1'028
$Q_{0.005} / Q_{0.995}$	<u>20.1</u>	24.6	21.3	21.9	1'451	-1'435
		Bootstrap of t-value with fixed weights				
$Q_{0.05} / Q_{0.95}$	1.53	1.71	1.38	1.52	1.49	1.31
$Q_{0.025} / Q_{0.975}$	1.72	2.00	1.63	1.83	1.75	1.55
$Q_{0.005} / Q_{0.995}$	<u>2.11</u>	2.31	1.94	2.18	2.20	2.12
		Bootstrap of t-value with stochastic weights				
$Q_{0.05} / Q_{0.95}$	1.22	1.56	1.17	1.31	1.24	1.38
$Q_{0.025} / Q_{0.975}$	1.36	1.78	1.44	1.54	1.47	1.61
$Q_{0.005} / Q_{0.995}$	<u>1.75</u>	2.19	1.94	2.11	2.28	2.05

Note: Estimation of stderror based on 10'000 bootstrap replications. Quantiles of estimate with fixed weights based on 10'000 replications. The other quantiles are estimated by double bootstrapping and only based on 1'000 replications in inner bootstrap and 1'000 replications in outer bootstrap. Aggregated LATE for labour markets 8, 9, 11, 21, 22, 23, 25.

Table D.3: Inverse of covariance matrix for aggregated treatment effect for compliers, Selection A

Outcome variable: Employment May-Aug. 1999							Outcome variable: Employment Sept.-Dec. 1999						
1.75	0.06	-0.06	-0.01	0.03	-0.10	-0.01	3.41	0.08	-0.07	0.09	0.02	-0.11	0.00
0.06	5.87	-0.06	-0.01	0.01	0.12	-0.04	0.08	7.53	-0.05	-0.04	-0.07	0.28	0.02
-0.06	-0.06	6.69	0.15	-0.07	0.09	0.03	-0.07	-0.05	4.68	0.21	-0.03	0.07	-0.01
-0.01	-0.01	0.15	14.93	0.02	-0.13	0.00	0.09	-0.04	0.21	17.37	0.08	-0.28	0.00
0.03	0.01	-0.07	0.02	7.12	-3.25	-0.17	0.02	-0.07	-0.03	0.08	4.13	-2.13	-0.07
-0.10	0.12	0.09	-0.13	-3.25	42.35	-0.14	-0.11	0.28	0.07	-0.28	-2.13	45.68	0.05
-0.01	-0.04	0.03	0.00	-0.17	-0.14	6.58	0.00	0.02	-0.01	0.00	-0.07	0.05	5.17
Outcome variable: High earnings May-Aug. 1999							Outcome variable: High earnings Sept.-Dec. 1999						
6.08	-0.06	-0.04	-0.11	0.02	-0.11	0.08	4.86	-0.04	-0.03	-0.17	-0.03	0.06	0.04
-0.06	6.22	-0.01	0.09	-0.03	0.06	-0.02	-0.04	6.70	-0.05	0.06	-0.02	0.10	0.02
-0.04	-0.01	8.77	0.18	-0.05	0.17	-0.04	-0.03	-0.05	7.16	0.15	-0.07	0.24	-0.02
-0.11	0.09	0.18	16.37	0.19	-0.37	-0.01	-0.17	0.06	0.15	15.81	0.13	-0.46	-0.13
0.02	-0.03	-0.05	0.19	6.69	-2.99	0.03	-0.03	-0.02	-0.07	0.13	5.11	-2.42	-0.01
-0.11	0.06	0.17	-0.37	-2.99	40.71	-0.12	0.06	0.10	0.24	-0.46	-2.42	38.55	0.05
0.08	-0.02	-0.04	-0.01	0.03	-0.12	6.50	0.04	0.02	-0.02	-0.13	-0.01	0.05	4.02
Outcome variable: Labour earnings May-Aug. 1999							Outcome variable: Labour earnings Sept.-Dec. 1999						
0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.26	0.00	0.01	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00
0.00	0.01	0.00	0.62	0.00	-0.02	0.00	0.00	0.00	0.00	0.62	0.00	-0.02	0.00
0.00	0.00	0.00	0.00	0.12	-0.07	0.00	0.00	0.00	0.00	0.00	0.07	-0.04	0.00
0.00	0.00	0.00	-0.02	-0.07	1.25	-0.01	0.00	0.00	0.00	-0.02	-0.04	1.22	0.00
0.00	0.00	0.00	0.00	0.00	-0.01	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.26

Note: Inverse of estimated covariance matrix of $\hat{\theta}$.