

DISCUSSION PAPER SERIES

IZA DP No. 10582

**The Impact of Traineeships on the
Employment of the Mentally Ill:
The Role of Partial Compliance**

Alberto Martini
Enrico Rettore
Gianpaolo Barbetta

FEBRUARY 2017

DISCUSSION PAPER SERIES

IZA DP No. 10582

The Impact of Traineeships on the Employment of the Mentally Ill: The Role of Partial Compliance

Alberto Martini

*Università del Piemonte Orientale
and ASVAPP*

Gianpaolo Barbetta

*Università Cattolica del Sacro Cuore di
Milano and Fondazione Cariplo*

Enrico Rettore

Università di Trento, FBK-IRVAPP and IZA

FEBRUARY 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

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 directly from the author.

ABSTRACT

The Impact of Traineeships on the Employment of the Mentally Ill: The Role of Partial Compliance*

Lavoro&Psiche is a RCT aimed at increasing employment among severely mentally ill patients, by offering them a structured job-search experience. The key feature of the treatment was the presence of a “job coach” entirely dedicated to support the job search of a small number (12-13) of mentally ill patients. What most often the job coach did was finding the patient a traineeship. If one were to consider only the effect of making the support of the coach available – the so-called Intention-To-Treat effect – it would be a disappointing statistically non significant 5 percentage point difference in the employment of treatment and control members, in the full post-treatment year. However, the impact of making something available is rarely the primary interest of policy-makers, who would rather know the effect of receiving it. The main difficulty in obtaining unbiased estimates of the latter is that ‘who receives what’ is no longer solely determined by randomization, but also by post-randomization events and decisions. During the implementation period of Lavoro&Psiche of 2011-12, an unprecedented wave of traineeship opportunities hit those enrolled in the demonstration, both in the experimental and in the control group. The main finding is that, for the subset of the experimental group that was induced by the offer to be involved in a traineeship the chances of having a job in the post-treatment year more than double from the 17% of those who did not do any traineeship to the 34% of those who did, well above the Intention-To-Treat impact estimates.

JEL Classification: J78, J48, J38

Keywords: mental illness, RCT, traineeships, ITT, tests for selection bias, placebo tests, instrumental variables

Corresponding author:

Enrico Rettore
Department of Sociology and Social Research
University of Trento
Via Verdi 26
38122 Trento
Italy
E-mail: enrico.rettore@unitn.it

* Many people contributed to the design and implementation of the Lavoro&Psiche demonstration. Fabio Dusio, Luigi Nava and Fabio Sandrolini from ASVAPP, Paolo Canino, Davide Invernizzi and Filippo Petrolati from Fondazione CARIPLO, Giovanni Ciniselli and Andrea Quarenghi from Ala Sacco. Moreover, we would also like to thank Stephen Bell and Barbara Romano for helpful comments on earlier drafts of the paper.

1. The dismal employment situation for the mentally ill

The extremely low labour force participation of people with severe mental illness (SMI) represents a serious public policy problem in most countries. Little consensus exists on what the right mix of policies ought to be to address this problem. Opinions vary widely, from those who still believe that the mentally ill should be confined to protected or sheltered environments to those advocating immediate placement in competitive employment, with some external support for job search and job retention. In addition to the sheltered/competitive jobs controversy, there are other dilemmas: one is to avoid the disincentives produced by disability insurance, the other is to fight discrimination by imposing hiring mandates on private employers (the first of these being on the policy agenda in the United States, while the second is more relevant for most of continental Europe, e.g., France, Germany, Spain and Italy), a rather direct and ultimately weak way to overcome the deep-rooted barriers faced by people with SMI.

Even compared with persons with physical or sensory disabilities, the mentally ill fare particularly badly. However, this gap is often very difficult to quantify, because rarely employment statistics are broken down by type of disability. There are attitudinal as well as structural barriers that prevent people with mental illness from becoming active participants in the competitive labour market. It is important to keep in mind that there aren't just economic benefits to employment. For those with severe mental illness, work can provide a sense of purpose, improve self esteem, and even lessen psychiatric symptoms (McGurk, et al., 2005). Rowland and Perkins (1988) identified four benefits of work: work as a restorative psychological process, work to improve self-concept, the protective effect of work and the social dimension of work. Positive and meaningful employment experiences have been linked to improved self-concept and self-efficacy, higher ratings of subjective well-being, regaining self-esteem, improved engagement in work activity with associated symptom reduction. Most importantly, work offers hope, which is vital to recovery from mental illness (Boardman et al. 2003). To be excluded from work erodes self-confidence and creates a sense of isolation and marginalization.

There are undoubtedly many barriers to work for people with SMI: barriers related to the symptoms of the disease themselves, such as cognitive impairments that can be found in attention deficit, psychomotor speed, verbal, learning and executive function (McGurk & Wykes, 2008). Though these impairments can create complications for any job, they are even more of a detriment in a highly technological world, where individuals with cognitive impairments often face idiosyncratic technology-related difficulties. (Guenther, 2013).

Other barriers originate in the workplace itself. Stigma associated with mental illness creates a reluctance to hire the mentally ill and foster low expectations of mentally ill workers in general. There is plenty of evidence pertaining to employment-related stigma and discrimination experienced by people with mental disabilities. Stigma is likely the single major cause of employment inequity for people with a mental disability who experience direct discrimination because of prejudicial attitudes from employers and workmates and indirect discrimination owing to historical patterns of disadvantage, structural disincentives against competitive employment and generalized policy neglect. But what perhaps singles out the stigmatization of people with mental health problems, from other potentially marginalized groups, such as women, ethnic minorities and people with physical disabilities, is the voice that they often do not have in fighting against discrimination. They are one of the most marginalized group in most societies.

Italy offers a good example of an ambitious though it seems largely ineffective legislation against discrimination for all persons with disabilities. A major reform was introduced almost 20 years ago, with Law 68 of 1999. This legislation strengthened a pre-existing system of mandatory hiring of persons with disabilities, not only mental but mostly physical and sensorial. It represents the evolution of the Italian legislation in terms of occupation of people with disabilities, while integrating it with the emerging principles of the international norms aimed at protecting the rights of persons with disabilities. Law 68/99 applies to people in working age with physical, mental or sensorial handicaps, with a reduction of their working ability of more than 45%, confirmed by ad hoc commissions for the certification of disability. According to Law 68, both private and public sector employers are required to hire a certain percentage of disabled workers, based on the size of their workforce:

- Employers with more than 50 employees must meet a 7% disability employment quota;
- At least 2 disabled workers must be hired in workplaces of 36 to 50 employees;
- Workplaces of 15 to 35 employees must hire at least 1 disabled worker if they operate new intake.

These provisions are built around the same notion of mandatory hiring, limited to new workers and valid for technical/executive staff only. Employers who do not meet the disability employment target must pay a fine to a specific fund. This fund is managed at the regional level and should work on furthering the integration of disabled people in the labour market. In 2013, close to 68.000 persons nationwide applied to join the “Law 68 lists”. We know that, in 2013, 18,300 persons were hired from the list. The ratio of hired from the list and people joining the list is perhaps a more telling figure: this number has been fairly stable despite a changing macroeconomic scenario at around 25%. So only one in four at the most might benefit directly from being from the “Law 68 list”, while the benefit to people with SMI might be substantially lower,

2. Enter Supported Employment

Supported employment is distinguished from other approaches like vocational rehabilitation by its emphasis on rapid job search (rather than extensive vocational assessment or training), placement in competitive jobs (rather than set-aside jobs for persons with disabilities), integration of vocational and clinical services, attention to individual preferences with respect to job types, disclosure of psychiatric condition to employers, and follow-along supports to facilitate job retention (McGurk, et al., 2005). Empirical evidence for this type of programmes is strong, with multiple randomized controlled trials demonstrating the superiority of supported employment over a variety of other rehabilitation models, including group skills training, sheltered workshops and psychosocial rehabilitation services. Such compelling evidence from controlled trials for supported employment started being gathered in the US around the turn of the century (Bond, et al. , 2002) and it is now widely accepted as an evidence-based practice for severe mental illness, and efforts are still under way to disseminate it (Bond, et al. , 2012).

As early as 2007, Lancet published a synthesis of the results of EQOLISE (Enhancing the Quality of Life of Individuals with Supported Employment), a multisite/multi-country RCT to test the effectiveness of the Individual Placement and Support (IPS) versus the more traditional vocational training approach to the placement of persons with SMI (Burns et al., 2007). Part of the international team that implemented EQOLISE was a group of Italian mental health clinicians that operated in a Northern Italian Region, Emilia Romagna, and the coastal city of Rimini was included in the study, along with London (UK), Ulm-Guenzburg (Germany), Zurich (Switzerland), Groningen (the Netherlands) and Sofia (Bulgaria). Patients were included if they had a diagnosis of SMI (psychotic illness including bipolar disorder). The EQOLISE study replicated

to some extent the excellent results obtained in US trials, despite the extensive differences in labour market regulations, organization and culture of mental health services. In this international six-site randomized trial, IPS was superior to the vocational approach for the number of people entering the competitive market, the number of days and hours worked, and the amount of money earned. The 312 patients with severe mental illness were randomly assigned in six European centres to receive IPS (n=156) or vocational services (n=156). Patients were supported for a maximum of 18 months. The primary outcome was the difference between the proportions of people entering competitive employment in the two groups (55% vs. 26%): these are intention to treat estimates.

When deciding which approach to take in implementing the Lavoro&Psiche demonstration, the CARIPLO Foundation considered the possibility of keeping it simple by replicating and testing the IPS treatment. This plan was eventually shelved since the foundation was at that time more interested in testing a specific policy innovation against a Treatment as Usual (TAU), rather than two alternative approaches, like supported employment vs. vocational training. With the benefit of hindsight, that decision appeared even more sensible as the deepening of the economic recession made the ‘rapid job search’ at the basis of the orthodox IPS look more and more like wishful thinking.

Table 1 contains three pairs of columns. The first pair refers to the results just quoted for the employment impact of EQUOLISE. The next two columns refer to the Lavoro&Psiche demonstration. We used data from another region in northern Italy, Piemonte, adjacent to Lombardia and for which we happened to have access to the full PES archives, to provide a benchmark for the data produced by the demonstration. In both cases we selected the 2010 cohort—that is, in Piemonte those recently identified as disabled who enrolled in the law 68 list during 2010--and in Lombardia all those who were recruited by the Lavoro&Psiche demonstration between the Fall of 2009 and December 2010.

TABLE 1: Percent WITH A JOB at any point during the year, various cohorts

	Enrolled in EQUOLISE before 2006 in 6 cities		Enrolled in Lavoro&Psiche in 2010 in Lombardia		Enrolled in Law 68 list in 2010 in Piemonte	
	Control group	Treatment group	Control group	Treatment group	Physically Disabled	Mentally ill
1 year	n.a.%	n.a.%	21%	21%	33%	10%
2 years	26%	55%	23%	27%	36%	17%
3 years	n.a.%	n.a.%	25%	30%	37%	15%
4 years	n.a.%	n.a.%	n.a.	n.a.	37%	18%

Source: own calculations using PES data, EQUOLISE and Lavoro&Psiche data n.a.= not available

The results shown in the last two columns in Table 1 are noteworthy because they compare directly the same employment definition for the mentally ill and for the physically disabled, thus showing that the claims of discrimination against the mentally ill: over a third of the physically disabled have a job at some point over the four year horizon after enrolling in the Law 68 list, while the corresponding percentage for the mentally ill remains below 20%.

The middle two columns of Table 1 refer both to mentally ill patient, and show what some would define as the central finding from the whole Lavoro&Psiche demonstration. During 2013 (the first post-demonstration year, the third from enrolment) 25% of control group patients had some paying job at some point during the year, versus 30% of the experimental group, not statistically significant at conventional

levels, and up from 21 percentage points for during the first year. Such 5-percentage point difference are just of the possible estimates of the impact of the treatment on the outcome. They are called intention-to-treat (ITT) and they do not measure the effect of actually receiving the treatment, rather that of being offered one. These estimates are reproduced below in Table 4, column (1).

However, the evidence points in a different direction, namely that of widespread non-compliance. The major of these pieces of evidence on non-compliance has to do with the role played by the traineeships, which are not given particular emphasis in the protocol but as a matter of fact had fundamental role in the implementation. Traineeships are not employment, although some trainee might be productive very soon. However, the motivations and expectations behind the hire of an employee are fundamentally different than those for a trainee. For many employers offering traineeship positions in an action of Social Corporate Responsibility (while for others is an act of normal.) Anyway, as shown in Table 2, the mentally-ill in Piemonte fare slightly better with respect to the physically disabled, most of whom had prior work experience. The striking difference is the one between the next two columns, taken from the Lavoro&Psiche demonstration in Lombardy, and their Piemonte benchmark.

TABLE 2: Percent in a TRAINEESHIP at any point during the year, members of the 2010 cohort

	Enrolled in Law 68 list in 2010 in Piemonte		Enrolled in Lavoro&Psiche in 2010 in Lombardia	
	Physically disabled	Mentally ill	Control group	Treatment group
After 1 year	4%	9%	27%	46%
After 2 years	3%	5%	23%	42%
After 3 years	1%	3%	18%	12%
After 4 years	2%	4%	n.a.	n.a.
Source: own calculations using PES data and L&P data			n.a.= not available	

While in Piemonte the percentage of traineeships targeted to the mentally ill covers between 5 and 10% of the cases, we can confront the totally anomalous figures of the Lavoro&Psiche experiment, which go over 40 % per year in the treated group.

3. The Lavoro & Psiche random assignment demonstration

The Lavoro& Psiche demonstration was conducted in Lombardy between 2009 and 2012, with generous funding from the CARIPLLO Foundation. The demonstration aimed at producing evidence on the effectiveness of a particular strategy for increasing gainful employment among severely mentally ill patients—that is, offering them intensive counselling and coaching by caseworkers with an unusually small caseload: each case-worker, named job coach, supported 12-13 patients for a period varying between 24 and 36 months. 311 persons with SMI were referred to by the mental health staff of the 29 local mental health centres in four provinces of Lombardy the end of 2010, of which 157 were randomly assigned to the experimental treatment, and 154 to the control group, whose members retained complete access to the existing services, but not to the unique feature of the demonstration, the support of a job coach. Table 3 contains full descriptive statistics for the two samples, information recovered from the baseline interview or from administrative archives.

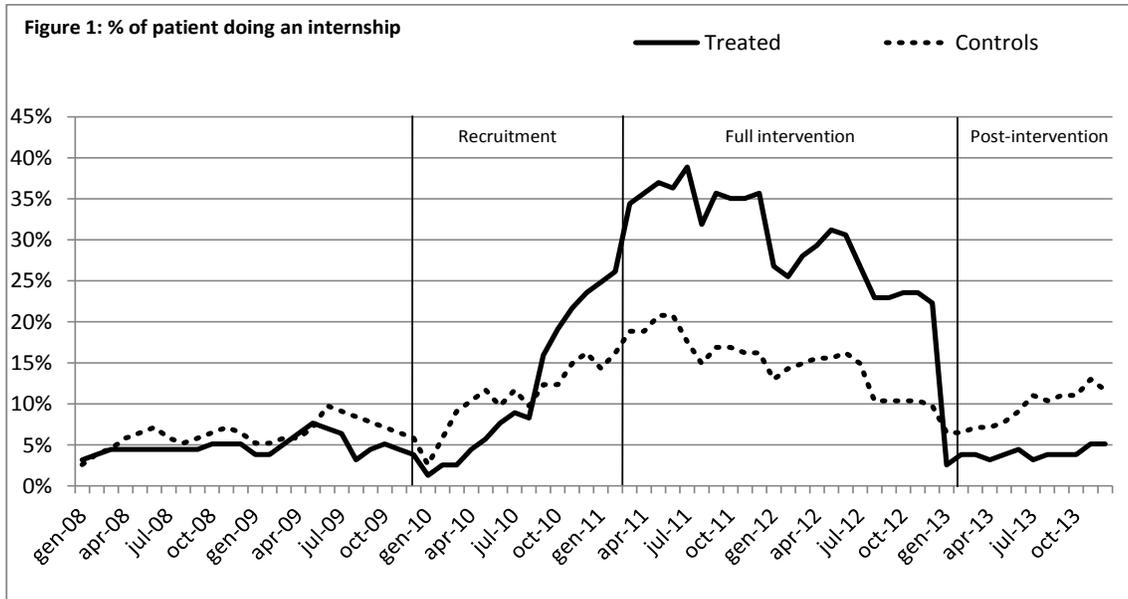
Table 3. Pre-treatment characteristics of the treated and control samples

	Treated	Controls	<i>p-value</i>
N.	157	154	
Gender			
Female	38%	38%	<i>0,8946</i>
Age			
<i>Mean</i>	<i>36,1</i>	<i>35,7</i>	<i>0,6570</i>
Age group			
Up to 24 years	8%	8%	
25 - 34	38%	33%	
35 - 44	41%	47%	
45 years or more	14%	11%	
Education			
Primary school	1%	1%	
Middle school	39%	35%	
Vocational school	17%	19%	
High school diploma	29%	33%	
University Degree	5%	5%	
Missing	8%	6%	
Diagnosis			
Bipolar disorder (F31 dell'ICD-10)	15%	14%	
Borderline personality disorder (F60.3 dell'ICD-10)	13%	12%	
Paranoid personality disorder (F60.0 dell'ICD-10)	10%	8%	
Schizoid personality disorder (F60.1 dell'ICD-10)	7%	5%	
Schizophrenia and related disorders (F20 dell'ICD-10)	55%	61%	
Disability percentage			
More than 45%	61%	57%	
<i>Average disability percentage</i>	<i>77,11</i>	<i>78,84</i>	<i>0,4166</i>
N.	95	88	
Previous work experience			
Had a job in 2008	38%	33%	<i>0,3496</i>
Had a job in 2009	28%	27%	<i>0,8825</i>
Done an internship in 2008	8%	11%	<i>0,4116</i>
Done an internship in 2009	10%	13%	<i>0,4426</i>
N.	157	154	

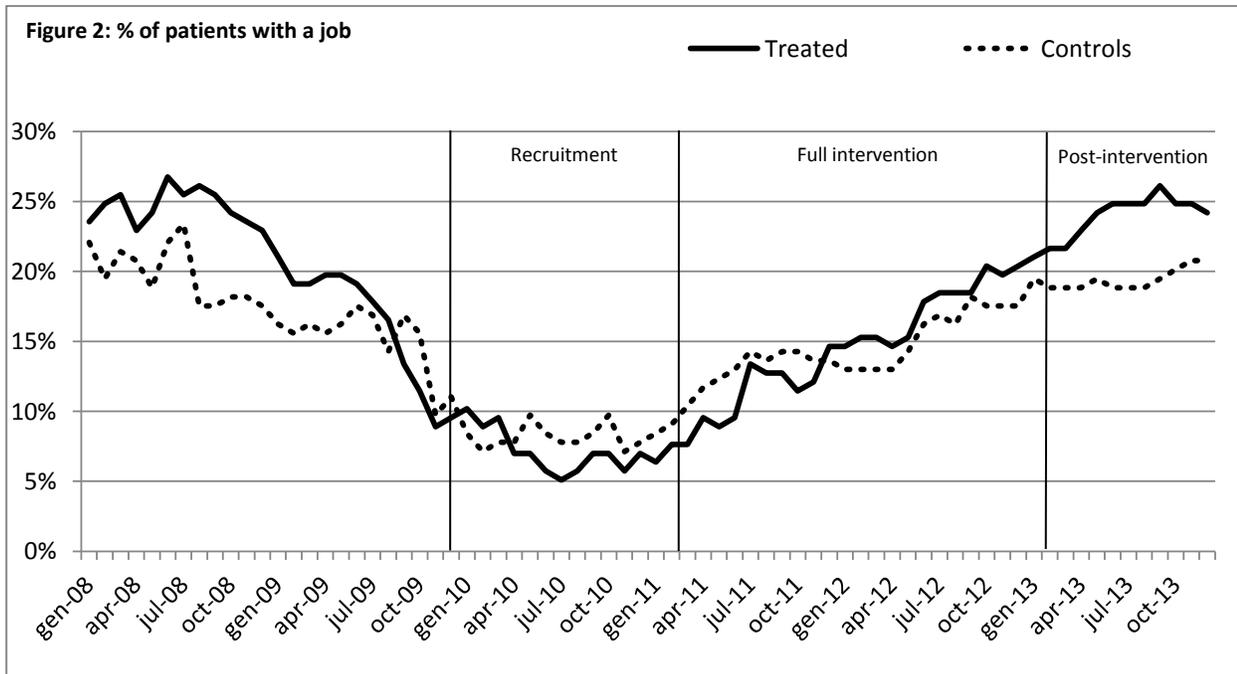
Source: Own calculation based on baseline interviews and COB data

The recruitment of patients started in late 2009 and continued throughout 2010. Upon referral to the project by the mental health center staff, patients went through a face-to-face baseline interview with an ASVAPP representative, their informed consent was obtained and then they were randomized into experimental and control groups and immediately notified by the center staff. Blocking was used to increase precision, by stratifying on gender, age, and location. For an average period of about two and half years, the control group had access to the job placement services normally available to all mental patients, while the treated group members received help in their job search by trained, qualified and committed job coaches, carrying a light workload of 12-13 patients.

The outcomes presented in this paper are limited to employment, a choice that might sound odd when taken with respect to individuals whose defining feature is a serious health condition. The reason for this choice is two-fold: on the one hand, employment remains the specific focus for the whole demonstration; on the other hand data availability admittedly carried some weight: around 2008, Italy made a huge leap in the collection of quality employment data, by creating a legal obligation for every employer to communicate to the local PES office information on all job contracts (in fact this is usually called the COB archive, which stands for ‘Comunicazioni Obbligatorie’.)



Using the COB data we were able to construct monthly individual work and training histories from 2008 through the end of 2013. We were able to distinguish between paid employment and traineeships. These two separate monthly histories are shown respectively in Figure 1, which shows the monthly figures for the fraction of the treated and controls who had a traineeship, while Figure 2 shows the same for paid employment.



A sharp difference in the time profiles for the two types of positions in the labor market is immediately noticeable. While the monthly figures for the traineeships closely parallel the various implementation phases of the demonstration, the monthly figures for employment ignore completely the pattern determined by the demonstration. Not even a minimal discontinuity is noticeable in the level of employment, not even in correspondence of the sharp phase-out of the demonstration: all job coaches were laid off in December 2012, a fact known to all involved since the Spring of that year. The evident sharp drop of over 20 percentage point between December 2012 and January 2013 means simply that the support of the coach is deemed necessary for the very existence of a traineeship. A sort of clause like “no traineeships without external support from a coach or somebody else.” To us, this evidence reinforces the notion that the traineeship should not be considered an outcome to be mixed with paid employment: rather, the best way of explaining the massive increase in traineeships is that, during the implementation of the demonstration, locating firms willing to offer traineeships to mentally ill patients became the major activity of the coaches, thus making it the essential active component of treatment.

While the members of the treatment group might have received less than what was planned in the slightly grandiose Lavoro&Psiche protocol, the members of the control group might have received more than envisioned in the design, due to a complex set of circumstances, which we try to explain as concisely as possible. First, a fundamental design choice was made, in favour of randomizing patients within sites (the 29 local mental health centres) rather than whole sites. The choice that was made had the advantage of minimizing the number of patients to be treated. To reach the same power for the test, randomizing sites requires to have more treated cases, thus making the study more expensive. On the other hand, randomizing patients within sites has the possible risk of invalidating the whole study, because of the possibility of contamination—that is, the members of the control group receiving services that might be formally different, but from a practical point of view might be so similar to dilute away any difference in the treatment received by treated and controls. Moreover, randomizing treatment within local health units, created considerable stress for the local health staff, who (i) saw their patients being assigned to either treated or controls without being consulted (which is the essence of random assignment, an idea that most

of those operating in the social sector loath, often with a vengeance) (ii) had to explain the patients the meaning of the informed consent and, last but not least, (iii) had the unpleasant duty of communicating to each patient their status with respect to receiving the support of the coach vs. reconfirming their access to the existing services. Given that most of the local health staff in charge of job placement (along with many other duties) had built through the years a network of firms willing to offer traineeships to mentally ill-patients, the temptation was there for using such resources to compensate for the “unjust” situation created by the randomization: for many such temptation was too hard to resist, and they caved in very easily. To this we might add that the Lavoro&Psiche program was met with some hostility, due to the fact the existing mental health staff had both much more experience than newly hired job coaches, but much less time available to devote to job placement. This latent conflict was, if anything reinforced by some of the choices made during the design phase and discussed above (such as the randomization taking place at the individual level, as well as placing the burden of explaining the informed consent and of notifying the outcome of randomization to each patient given to hostile health operators).

Summarizing, the inclusion of Lavoro&Psiche in the fairly structured mental health system of the Lombardia region created tensions along some predictable fault lines. As in most other cases of non compliance, the trouble for reaching credible causal conclusions is that differences in behaviour taking place after randomization are de facto trying to undo what randomization does. While in our case the “classical” cross-overs are not possible, because the coach are not allowed to make substitution or addition to the list of patients they receive for coaching. If the active treatment tends to reduce to offering a traineeship, it becomes much easier for the controls to receive some treatment, because nobody is expected to receive much more than a traineeship. We call this a case of inverted cross-overs, in which it is the treated who get less, rather than the controls managing to get more.

Let us look again at Figure 1 above: while back in 2008-09, before the demonstration started, the monthly rate of utilization of traineeships was about 5% of both groups of patients later to become treated and controls. In 2011-12 we get an unduplicated total of 71% and 43%, respectively, of patients in a traineeship. As far as non compliance goes, this situation is hard to beat. What more realistically has been implemented with Lavoro&Psiche is a *randomized encouragement trial*, in which patients are only randomly assigned to two different levels of encouragement: those assigned to the coach receive a higher dose of encouragement, by a job coach; and the controls, who never see a coach, are stuck with whatever the mental health staff is willing to find for them.

The main implication of this story is that, for the subset of experimental that were induced by the offer to be involved in a traineeship and thus complied with the treatment assignment, the impact of Lavoro&Psiche is larger than it is for everybody else. As explained in more detail below, our best estimate of the chances of holding a job it almost doubles from 17% among those who did not do any traineeship up to 34% for those who did. In order to explain this result, we need a slightly more analytical apparatus. Technically speaking, we are interested in treatment-effect variation across partially or fully latent subgroups defined by post-treatment characteristics.

It is useful to introduce some formal notation. Let (L_1, L_0) be potential outcomes, where L_1 is the employment status in 2013 in case of taking part in a traineeship during the experimental period and L_0 is the employment status in 2013 in case of not taking part in a traineeship during the experimental period.

The causal effect of taking part in a traineeship on the employment status in 2013 at the individual level is $\beta = L_1 - L_0$. Following the argument developed in the previous section this is our target causal parameter.

The treatment status T is equal to one for those taking part into a traineeship during the experimental period, it is equal to zero for those who did not. This treatment status is not randomly assigned, it is the job coach to be randomly assigned. But as a matter of fact being assigned a job coach makes a major difference for the probability of taking part into a traineeship (see Fig. 1). Let Z be the randomly determined assignment of patients to a job coach. There is plenty of room for two-sided non compliance, since there are both no-shows -- i. e. patients assigned a job coach and not taking part in any traineeship ($Z=1, T=0$) -- and cross-overs -- i.e. patients not assigned to a job coach and nonetheless taking part in a similar traineeship ($Z=0; T=1$). It should be noted that all traineeships are born equal, in the sense that they are all regulated by the same national law. Moreover, the *Lavoro&Psiche* is a tiny affair compared to the size of the labor market: so, holding constant the characteristics of the trainee, the traineeship experience was pretty much the same kind of experience across the two samples.

In one of the most influential papers of the last quarter century, Angrist, Imbens and Rubin (AIR,1996) showed that using Z as an instrumental variable for T allows to recover the so called Local Average Treatment Effect (LATE). It is local in that it is the (average) causal effect for the subpopulations whose treatment status T is affected by the value of the instrumental variable Z , the compliers in the AIR terminology. The key condition for the validity of this strategy is that being assigned a job coach must not have any effect on the outcome but the one through its effect on the probability of experiencing a traineeship spell¹. This is the so called exclusion restriction on which any IV strategy rely. This set-up might seem to be a post-hoc rationalization driven by data availability. Considering as the treatment only something we do observe -- namely the traineeship status T -- in principle might mean ignoring the existence of other, possibly beneficial, services provided as components of the treatment. The anecdotal evidence we gathered suggested otherwise. Thus we are convinced of its plausibility. Moreover, here below we show that the exclusion restriction bears testable implications.

When both the treatment status T and the instrument Z are binary variables, the IV estimator is the so-called Wald ratio, in which the causal effect on the outcome of being assigned a job coach is divided by the compliance rate, i.e. the difference in the take up rate between those assigned and not assigned, respectively, a job coach:

$$(2) \quad \text{LATE} = \frac{P(L = 1 | Z = 1) - P(L = 1 | Z = 0)}{P(T = 1 | Z = 1) - P(T = 1 | Z = 0)}$$

As applied to the case of *Lavoro&Psiche*, the LATE is as large as 18.9 points (see Table 4, column 2) more than three times the 5 points we got from estimating the causal effect on L of being assigned a job coach. At the close of the demonstration, over 71% of the experimental had been in a traineeship, versus 43% of the controls.

¹ There is also an additional technical condition named monotonicity: it must not be the case that some patients taking part into a traineeship if not assigned a job coach would not take part into the traineeship in case they are assigned a job coach.

There are a couple of problems with this estimate. First, it is fairly imprecise (see the standard error in col. 2 of Table 4), which is typically the case in most IV applications (Murray, 2006).

Table 4. Comparing Estimates of the Employment Effect of a Traineeship

	ITT est.	IV est.	OLS est. does not include Z	OLS est. include Z	OLS for those in traineeship	OLS for those not in traineeship	OLS on 2008 outcome	OLS on 2009 outcome
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.251 (.035)	0.165 (.105)	0.166 (.0381)	0.166 (0.042)	0.343 (.059)	0.172 (.042)	0.356 (.0418)	0.273 (.0391)
Offered a traineeship	0.053 (.051)	-	-	0.0011 (.0518)	0.014 (.074)	-0.017 (.069)		
Took a traineeship	-	0.189 (.177)	0.185 (.0502)	0.185 (.0524)			0.0015 (.0551)	0.0066 (.0515)
N	311	311	311	311	179	132	311	311

(Standard errors in parentheses)

Second, AIR show that the denominator of the LATE is an estimate of the size of the sub-population of compliers. In this instance it is as large as 28% of the whole sample. Strictly speaking the estimated average causal effect refers only to this specific sub-population. What is worse, who exactly the compliers are is not known raising a problem of external validity for this estimate.

In a recent paper Black *et al.* (2015) propose a simple test for the presence of selection bias in an IV set-up. On accepting the null hypothesis – i.e. no selection bias – one can use the (typically much) more precise OLS to estimate the causal effect of T on L instead of using the IV. Moreover, on accepting the null, one gets evidence consistent with the case in which the LATE coincides with the average causal effect over the whole population (ATE).

The test amounts to run the OLS regression of the outcome L on the instrumental variable Z separately for the two sub-groups indexed by T. Consider the case of T=1. All these patients went through a traineeship. Those with Z=0 got it despite having been randomly denied the job coach. In the AIR terminology they are the always takers. Those with Z=1 got it either because they have been assigned a job coach – the compliers – or because they are always takers. Finding that the potential outcome L_1 does not depend (on average) on Z means that always takers and compliers are equivalent (on average) with respect to L_1 . In the case of the sub-group T=0 the story is the same with the always takers replaced by the never takers, those who do not undergo a traineeship whether assigned or not a job coach.

Black et al. (2015) also note this test has power against the violation of the exclusion restriction. This is because if the treatment status Z had an impact on the outcome via any channel other than the traineeship status T then conditioning on T the outcome would depend on Z.

Table 4, columns 5 and 6, shows that the null hypothesis is not rejected for both sub-groups. Even pooling together the two groups to gain precision in the estimation of the coefficient on Z the results do not change (see column 4), the estimated value of the coefficient is both tiny and statistically insignificant. Summing up, no evidence of selection bias nor of violation of the exclusion restriction.

To provide a further check that selection into traineeship is *not* driven by individual characteristic relevant for the probability to get a job we ran the OLS regression of the *pre-treatment employment status* on the traineeship status. We define the pre-treatment employment status the same way as the main outcome of our analysis but with reference to the calendar years 2008 and 2009, respectively. Finding that the employment status in either of these years depends on the traineeship status T in years 2010 to 2012 would be a clear sign that individuals experiencing a training spell are different from those who do not with respect to characteristics relevant for their employment status. Results in columns 7 and 8 show that no such effect exists. Reassuringly, there is no evidence pointing to self-selection into traineeship driven by unobservables relevant for the probability to get a regular job.

Based on the result of this set of specification tests we conclude that the OLS estimate in Table 4 column 3 deserves a causal interpretation. There is a major gain of precision switching from the IV to the OLS estimator (compare columns 2 and 3): we safely conclude that having done a traineeship doubles the probability of having a paying job a year later. Moreover, following Black et al. (2015) the evidence that controlling for the traineeship status the assignment to the treatment does not matter for the outcome lends support to our claim that there was no service but traineeship provided by the treatment. Finally, the evidence that compliers are equivalent on the one hand to always takers (with respect to L_1) on the other hand to never takers (with respect to L_0) lends support to the external validity of this estimate.

6. More experimental evidence coming (not too) soon

Despite the uncertainty surrounding the approach discussed in this paper, and to some extent stimulated by it, one of the original partners of the Lavoro&Psiche demonstration, in partnership this time with the regional government of neighboring Piemonte, successfully bid for one of the grants under the European Union initiative on social innovation through experimentation (PROGRESS EaSI). The project acronym is TSUNAMI (*Traineeship as a Springboard out of UNemployment for those Affected by Mental Illness*): the role of the coach is set explicitly as 'scouting for traineeships', while finding paid employment remains in the background as a more distant goal which it might not be worth measuring too soon. This new experiment would shed light on whether providing a structured training opportunity would induce those at the margin to take such an opportunity, which would have (testable) consequences for their entering or not the labor market. Unfortunately it will take three full years to obtain the complete results on the effects of devoting resources in this type of programs. To be sure, it is unlikely that the dismal employment situation for the mentally ill will be solved in the next 3 years.

Another contribution from such sequence of experiments is the opportunity it offers of making the case for a wider adoption of randomized control trials—even in a policy-making context more oriented toward reaching spending targets and less inclined to asking questions on what truly works and for whom, than it is the case in the US context. The authors, while being fairly sceptical that any RCT-mania will take Europe by storm any time soon, also recognize that any possibility would require the informed and coordinated effort of the different institutions involved on a given policy issue, from national and local governments to philanthropic foundations, from policy research organizations to the media.

References

- Angrist, Joshua, G. Imbens, D. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables.", *Journal of the American Statistical Association* 91(434): 444-55.
- Black Dan, Joonhwi Joo, Robert LaLonde, Jeffrey A. Smith, Evan J. Taylor, (2015), "Simple Tests for Selection Bias: Learning More from Instrumental Variables", IZA DP 9346, September 2015
- Boardman, J. Grove, B. Perkins L. Shepherd G.M J (2003) "Work And Employment For People With Psychiatric Disabilities" *British Journal of Psychiatry*, 210 (2) A7
- Imbens, Guido and Don Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction* (2015), Cambridge University Press, UK ,1-650.
- Bond G, Drake RE, Becker DR (2012) Generalizability of the Individual Placement and Support (IPS) model of supported employment outside the US. *World Psychiatry*. 11:32-39.
- Burns T, Catty J, Becker T, Drake RE, Fioritti A, Knapp M, Lauber C, Rossler W, Tomov T, van Busschbauch J, Whita S, Wiersma D (the EQOLISE Group. (2007). The effectiveness of supported employment for people with severe mental illness: A randomized controlled trial in six European countries. *The Lancet*, 370, 1146– 1152.
- McGurk, Susan R., Mueser, Kim T., Pascaris, Alysia. (2005). "Cognitive training and supported employment for persons with severe mental illness." *Schizophrenia Bulletin*, 898-909.
- Murray, Michael P. (2006). "Avoiding Invalid Instruments and Coping with Weak Instruments," *Journal of Economic Perspectives*, American Economic Association, vol. 20(4), pages 111-132
- Rowland L. and Perkins (1988) 'You can't eat, drink or make love eight hours a day'. The value of work in psychiatry. *Health Trends*, 20, 75—79.