THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 12/13

Does Coordination of Welfare Services' Delivery Make a Difference for Extremely Disadvantaged Jobseekers? Evidence from the 'YP^4' Trial

Jeff Borland, Yi-Ping Tseng and Roger Wilkins

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Does Coordination of Welfare Services' Delivery Make a Difference for Extremely Disadvantaged Jobseekers? Evidence from the 'YP[4]' Trial*

Jeff Borland[†], Yi-Ping Tseng[‡] and Roger Wilkins[‡]

[†] Department of Economics, The University of Melbourne

[‡] Melbourne Institute of Applied Economic and Social Research, The University of Melbourne

Melbourne Institute Working Paper No. 12/13

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-4301-6

March 2013

* This research reported on in this paper was supported by an Australian Research Council Linkage Grant (LP0668208) and by cash and in-kind contributions from Hanover Welfare, Brotherhood of St. Laurence, Melbourne Citymission and Loddon Mallee Housing Services. We also acknowledge the Department of Families, Housing, Community Services and Indigenous Affairs and Centrelink for facilitating implementation of the YP[4] trial, and the Department of Education, Employment and Workplace Relations for providing administrative data. Thanks to David Black for research assistance, the YP[4] team at Hanover Welfare for compiling the data, and Louise Coventry and John Haisken-DeNew for comments on earlier versions of this paper.

Melbourne Institute of Applied Economic and Social Research

The University of Melbourne

Victoria 3010 Australia

***Telephone* (03) 8344 2100**

***Fax* (03) 8344 2111**

***Email* melb-inst@unimelb.edu.au**

***WWW Address* http://www.melbourneinstitute.com**

Abstract

Fragmented welfare service delivery has been identified as a significant barrier to improving outcomes for highly disadvantaged individuals. The 'YP[4]' trial, conducted from 2005 to 2009, sought to evaluate, by randomised control method, an approach proposed by Campbell et al. (2003) for integrating delivery of employment, housing, health and other services for young homeless jobseekers. Rather than providing extra access to services or utilisation of different services, the YP[4] trial involved assignment of a case manager to help tailor and coordinate available services to reflect the specific circumstances of young homeless jobseekers. We find that the YP[4] program did not have a significant effect on economic or psychological well-being, a finding that is robust to application of experimental and quasi-experimental methods. It is argued that our study contributes to knowledge on program design, particularly in relation to the importance of the scale of intervention and program administration.

JEL classification: I38, J08, J64, D04

Keywords: Welfare programs, disadvantaged jobseekers, randomised controlled trials, program evaluation

1. Introduction

A long-standing concern in implementation of welfare policy has been coordination of service delivery to individuals who have multiple sources of disadvantage. Perri (2004, p.1) has described the problem of improving coordination of policies as 'an eternal one' that dates back to the Chadwickian reforms of British local government in the 1830s. In Australia, as Campbell *et al.* (2003) and Parkinson and Horn (2002), among others, have discussed, a lack of overall coherence in service delivery is a problem acutely felt by government and community sector workers delivering assistance to young homeless jobseekers.

Integration of service delivery is intended to facilitate a holistic approach that takes account of the full circumstances of the individual and the full range of services available (Fine et al., 2005). This can allow tailoring of services to specific circumstances and, more particularly, for a coordinated approach to improving outcomes for individuals with multiple sources of disadvantage, where addressing all sources of disadvantage simultaneously is necessary to improve outcomes. In the case of homeless jobseekers, without addressing the problem of stable housing, it may be difficult to provide the basis for stable employment; but equally, the absence of stable employment may make it impossible to achieve a stable housing outcome. Implementing approaches that allow multiple sources of disadvantage to be addressed simultaneously has been widely advocated. Ways of achieving a more integrated approach that are cost effective and administratively workable are thus keenly sought by service providers.

'YP^4' was proposed as one possible model for improved coordination of service delivery, providing 'joined-up' services for young homeless job seekers in Victoria. The trial of YP^4 was conducted from 2005 to 2009 by a partnership of four community sector organisations with the cooperation of relevant government agencies.[1] Joined-up service provision in YP^4 involved the integrated delivery of employment, housing, health, educational and personal support services. The trial involved the (offer of) provision of joined-up services for between 18 to 30 months. Each participant was assigned a case manager who could provide advice and liaise with—and to some extent coordinate—service providers. The main objectives of YP^4 were to improve assistance to participants to find employment, provide secure housing tenure, improve health outcomes and increase engagement with local communities (Horn, 2004).

[1] The name of the trial derives from its target group and aims. 'YP' represents young people. The '4' refers to the four 'p's': purpose (a job); place (a home); personal support (offered by the project); and proof (evaluation).

The YP[4] trial was implemented as a random experiment. In each of four geographic locations (Cheltenham, Frankston, Bendigo and Inner Melbourne), 18 to 35 year-old homeless jobseekers eligible to participate were identified, and then randomly assigned to treatment and control groups, respectively known as J-group and S-group.[2] Recruitment into the trial took place over the period January to December 2005, all case management services ceased in June 2007, and final data collection was completed in early 2009.

In this paper, we describe this empirical analysis of the YP[4] trial, and present estimates of program impacts. The main effects of the YP[4] trial that we examine are on employment, welfare receipt, housing status, and health/well-being, as well as the amount and types of services accessed by participants. These program impact estimates are of outcomes under joined-up service delivery compared with outcomes under the status quo. That is, the treatment effect is not receipt of or access to particular services per se; rather it is receiving joined-up service delivery instead of standard service delivery. We focus on estimates of impacts on outcomes rather than attempting to evaluate the net social benefit of the program because little information on program costs is available and because many impacts are difficult to quantify in monetary terms, particularly those that relate to the well-being of the participant. For example, valuing psychic benefits of employment, stable housing or improved health is inherently difficult.

Our study is notable in several important ways—in being a randomised control trial; in its analysis of an intervention targeted at homeless jobseekers; in considering a new model for case management, YP[4], that sought to provide a more comprehensive approach to coordinating service delivery; and in the breadth of the outcome measures and time horizon that we use to examine the impact of YP[4].

First, there is only a handful of randomised control trials of social programs that have been undertaken in Australia. In the early 2000s, the Commonwealth Department of Family and Community Services undertook several randomised control trials to test the effect of interventions intended to increase information to sole parents on the availability of training programs (Barrett and Cobb-Clark, 2000), and to promote education and social participation by the very long-term unemployed (Breunig et al., 2003). At around the same time there was also a randomised control trial of the Drug Court system in NSW for dealing with persons with drug addictions and charged with criminal offences (Lind et al., 2002). This is the extent of application of the randomised control method to evaluation of social programs that we are aware of. Hence, we believe our study provides an important demonstration of the potential for applying this method of policy analysis, especially in

[2] However, as explained in Section 2, some J-group places were 'reserved' for assignment at the discretion of program service providers.

partnership with not-for-profit organisations. This is notwithstanding problems in the implementation of the randomised control trial methodology in this study, which we describe later.

A second notable feature of our study is its analysis of a new type of intervention, the model of joined up services. This model involves coordination across a broader range of service providers than in previous studies (beyond health care and housing providers), targeted at a population group, young homeless jobseekers, who have not been the subject of extensive previous research.

Internationally, there have been many studies of effects of health care and housing strategies for homeless people (see Fitzpatrick-Lewis et al. (2011) for a review of many of these studies), but relatively few randomised trials evaluating effects on employment and welfare dependence, or evaluating effects of more broadly-based interventions targeting the homeless. Interventions for homeless persons that have involved integrated or joined-up service provision in most cases have been accompanied by a new service or treatment, so that the distinct effect of 'joining up' services is not identified. For example, Toro et al. (1997), Bradford et al. (2005), McHugo et al. (2004) and Calsyn et al. (1998) all examine interventions involving some degree of coordination of service delivery to homeless people, but all involve a new treatment—which is the main focus of the studies. In the few studies focused on the effects of service delivery coordination, such as Rosenheck at al. (2002), the interventions have been restricted to health care and/or housing services, and in particular did not entail coordination with employment services providers. In addition, while there are many randomised trials and quasi-experimental studies of interventions to assist jobseekers, none of these appear to have been specifically targeted at the homeless. A limited literature on the effects of case management on outcomes for jobseekers (Lechner and Smith, 2007, and Berger et al., 2000), and a somewhat larger literature on effects of job counselling schemes (see Kluve, 2006 and Borland and Tseng, 2007 for surveys), do however exist.

A third feature of our study is the broad range of outcome measures we examine, and the time horizon over which we are able to test for effects of YP[4]. Through administrative data obtained from the Commonwealth government, and surveys undertaken as part of the YP[4] trial, we are able to measure outcomes including income support payment receipt, labour market outcomes, housing status, health and well-being, and community participation. These outcomes can be examined for up to three years after commencement in the YP[4] trial. This seems important given evidence that program effects may vary over time—due to factors such as accumulation of skills gained through a training program or the gradual build-up of cultural effects on behaviour due to program participation (see for example, Card et al., 2009, and Wolfers, 2006).

The remainder of the paper proceeds as follows. In Section 2 we describe the implementation of the YP[4] trial and the problems confronted. Section 3 describes the data available for our evaluation and the outcome measures that we derive from this data. In Section 4 we present estimates of the impact of YP[4] using alternative approaches: raw data on average outcomes, and a quasi-experimental matching method. On the basis that impacts are likely to be greater for participants who received more 'joined-up service provision', we also present estimates of the impact of YP[4] distinguished by duration of participation in case management. Concluding comments are provided in Section 5.

2. Design and implementation of the YP[4] trial

2.1 Design

The treatment group in the YP[4] trial was to receive joined-up case management, while the control group would receive standard service delivery. Random assignment of participants between treatment and control groups implied that the impact of 'treatment'—access to joined-up case management—would be able to be identified by comparing mean outcomes between the groups.

To be eligible for participation in the YP[4] trial, an individual was required to be aged 18 to 35 years, in receipt of Newstart Allowance or Youth Allowance (other), homeless or with a history of homelessness, and 'disadvantaged', as evidenced by eligibility for the *Personal Support Program* (PSP), *Job Placement, Employment and Training* (JPET) program or *Intensive Support-Customised Assistance* (ISCA).

The initial plan in the trial was for 130 homeless jobseekers to be recruited at each of four sites. At each site, individuals would be randomly assigned to treatment and control groups in a ratio of 60 treatment members to 70 control members. The recruitment process would therefore produce a total of 520 participants, comprising 240 treatment group members and 280 control group members. The higher rate of recruitment to the control group was motivated by an anticipated higher attrition rate.

Four partner agencies were involved in the trial, each being responsible for one site that broadly corresponded to the geographic area that agency primarily serviced. The responsible agency provided joined-up case management for all treatment group members recruited at its site. Management was provided by the Brotherhood of St. Laurence at the Frankston site, by Hanover Welfare at the Cheltenham site, by Melbourne Citymission at the Inner Melbourne site, and by Loddon Mallee Housing at the Bendigo site.

2.2 The YP[4] intervention

The key defining characteristic of joined-up case management was the assignment of a case manager to each treatment group member. Each case manager was an employee of one of the four partner welfare agencies and was tasked with meeting with the treatment group member on a regular basis, discussing their needs and concerns, directing them to appropriate services and, if required, assisting them to access those services. Case managers were also expected to liaise and foster relationships with service providers to facilitate better tailoring of services to participants' needs. Control group members were not assigned to a case manager, but could in principle access any of the services provided by the partner agencies, Centrelink, or Job Network (government-funded employment services mostly provided by non-government organisations), that were available to treatment group members. Indeed, in some cases control group members were required as a condition of income support payment eligibility to use services. It is therefore important to emphasise again that the YP[4] trial was not of a particular service or program, but a new integrated model of delivery of existing services.

The number of contacts that participants had with their case manager over the first two years of the YP[4] trial averaged 23, which approximately translates to a fortnightly frequency. However, the extent of contact varied considerably across participants, with 20 per cent having no contact, 17 per cent having 1 to 5 contacts, 19 per cent having 6 to 20 contacts, 21 per cent having 21 to 40 contacts and 22 per cent having 41 to 156 contacts (Grace and Gill, 2008). At any one time during the trial, there were between six and eight case managers employed across the four sites, suggesting that case managers were allocated relatively little time for each treatment group member. Perhaps as a result, the effect of the YP[4] program on service utilisation was limited. The results presented in Section 4 show no significant impacts of case management on utilisation of employment services or difficulties accessing services; and it appears to have reduced the number of services accessed in the first year after trial entry – although this latter finding has an ambiguous interpretation: it could reflect lower use of needed services or better matching of services to individual need. Overall, we believe that it is appropriate to interpret the YP[4] program, in the way it was implemented, as a relatively minimal intervention.

2.3 Implementation problems

Several aspects of implementation of the YP[4] trial appear to have compromised randomisation. Most importantly, assignment to treatment in the YP[4] trial was not completely random. Partly, this was by design, in response to partner agency requests for 'reserved' treatment group places that could be assigned to individuals at their discretion. Fifteen treatment group places were thus reserved in each

region for '…persons who happened to present to the YP[4] provider or another local service and seemed eligible' (Coventry and Pedrotti, 2006, p.6). Fifty-three of these places were ultimately filled. Unintended compromises of randomisation also occurred. Where this happened it was as a result of imperfect randomisation methods (such as assigning to treatment and control groups according to the day of the week) and asymmetric efforts in recruiting treatment and control group members (for example, making greater efforts to establish contact with persons assigned to the treatment group) (Coventry and Pedrotti, 2006).[3]

Implementation also suffered from difficulties in recruiting a sufficient number of participants. Only 445 participants were ultimately recruited, 75 short of the target number. Furthermore, achieving this number required changes over time in the methods used to identify potential participants. Thus, there are concerns that those participants recruited later are in some way different to those recruited earlier. Initially, recruitment was via 'spontaneous recognition', whereby Centrelink officers identified customers who came into the Centrelink Customer Service Centre (CSC) and appeared to meet the eligibility criteria for participation. This approach resulted in very few recruits, which led to the use of information available in Centrelink administrative records to identify potential participants. It is not possible to precisely identify persons eligible for participation using administrative data because homeless status is not explicitly recorded. Consequently, the criteria applied to select potential participants from the administrative data changed over time as refinements were made in an effort to better identify homeless persons. Both the move from 'spontaneous recognition' to searches of administrative data, and the refinement of the search criteria used in the data searches, had the potential to alter the composition of participants over time. The recruitment difficulties to some extent stemmed from the very specific criteria for participation in the trial. The requirement that participants be in receipt of unemployment benefits was particularly restrictive, excluding, for example, most single parents and people with disabilities.

Implementation and ultimately evaluation of the program was also adversely affected by the geographic specificity of recruitment and service provision. Some treatment-group participants who moved out of the area where they were recruited to the trial were not able to continue case management (although efforts were made to continue case management even if a participant had moved a considerable distance); and likewise, data collection was truncated for some treatment and

[3] Coventry and Pedrotti (2006) also note that there were differences in the implementation of the trial across the four sites, both in the recruitment of participants and in the nature of the case management services provided. For example, with regard to differences in recruitment, each site initially gave priority to recruiting a particular sub-group: in Cheltenham, priority was given to families, and no single persons under 25 years were to be recruited; in Inner Melbourne, no-one over 25 years of age was to be recruited; in Bendigo, priority was given to indigenous persons; and in Frankston, the option to identify a priority group was not taken up. Note, however, that these differences did not of themselves introduce non-randomness of assignment to treatment.

control group members who moved. Furthermore, Centrelink administrative data ceased to be collected if the participant was no longer resident in Victoria (although if the participant subsequently returned to Victoria, data collection would resume).

3. Data

3.1 Main data sources

Several administrative data sources were available for the evaluation of the YP[4] trial, supplemented by periodic surveys of treatment and control group members undertaken by Centrelink staff. These data sources are summarised in Table 1. These are the YP[4] recruitment data (1), YP[4] interview data (2), Centrelink payments administration data (3), Department of Education, Employment and Workplace Relations (DEEWR) employment assistance administrative data (4), and YP[4] case manager quarterly returns (5).

The YP[4] recruitment data (1) provides basic demographic information. This information has been used to evaluate the comparability of the treatment and control groups; and as a basis for matching treatment and control group observations in the program impact analysis.

Table 1: Data sources available to the YP[4] trial

Data Source	Information collected
1. Recruitment and referral data	Mostly demographic information collected at time of recruitment into the trial: region, sex, age, educational attainment, date of entry to YP[4], reserve place indicator, date of exit from YP[4], ex-offender status, indigenous status, and indicators for ISCA, PSP and JPET eligibility. Available for all participants.
2. Centrelink interviews	Interviews conducted by Centrelink staff annually (including at entry into YP[4]), supplemented with telephone interviews conducted by YP[4] staff with trial participants who have exited income support. This source provides information on accommodation (types used, number of moves, suitability, cost, financial support received), employment (types held, type seeking, barriers, activities undertaken to increase employability), educational attainment, use of community services, health and wellbeing, and community engagement. The supplementary interviews additionally collected information on reason for exit from income support and rate of pay.
3. Centrelink administrative data	Derived from payments' administration data. Provides details on benefit receipt (duration on each benefit type, benefit income, activities undertaken including periods of incapacity, reason for exit for each period off benefits), accommodation circumstances (type, rent paid and length of time at each location) and income from each employer while on benefit. Miscellaneous other information (participation in literacy & numeracy training, career counselling or voluntary work, full-time student status, breaches and suspensions) is also provided. Data is delivered annually.
4. DEEWR administrative data	Annual information on program expenditure, Job Network placement outcomes and program participation for each trial participant.
5. Case manager quarterly returns	Data provided by case managers on treatment-group members.

YP[4] interview data (2) and Centrelink administrative data (3) provide the main information on outcomes used in the assessment of the impact of YP[4]. The Centrelink administrative dataset is an annual extraction from Commonwealth government databases on income support (welfare) payment recipients. It contains individual-level information for the relevant year on duration on each benefit type and value of payments received for each benefit type. The YP[4] interview data consist of responses to an annually administered questionnaire on employment, income, housing, health, social and community participation, use of services and general wellbeing. This was collected by Centrelink officers for those on benefits and by YP[4] evaluators by telephone interview for those who had left benefits. Centrelink administrative and YP[4] interview data are available for three years following entry by a treatment or control group member to the trial.[4]

The Centrelink administrative data are close to complete for all treatment and control group members, whereas response rates are far short of 100 per cent for the YP[4] data. The response rate was 67.3 per cent for Year 1 interviews, 65.5 per cent for Year 2 and 38.3 per cent for Year 3. Due to the particularly low response rate in Year 3, for outcomes based on the YP[4] interview data, only Year 1 and Year 2 are examined. The minimal sample attrition in the Centrelink administrative data is an important feature of the trial that warrants highlighting, given the great difficulties normally faced in retaining such an itinerant group as homeless people.

While the response rate associated with the Centrelink administrative data is an advantage, the data also have limitations. Data are only available when individuals are on income support. First, the scope of outcome variables that can be obtained from the Centrelink data is limited. For this reason it was necessary to use the YP[4] interview data to obtain measures of outcomes such as health and wellbeing, and community engagement. Second, although the Centrelink data do include measures of earnings and moving house, two outcomes of considerable interest to assessments of YP[4], this is not available where an individual moves off income support payments. This is why we also undertake analysis utilising the YP[4] interview data on the restricted sample for which this data is available, since the interviews collected information on employment and housing status while off income support.

The other source of information used to evaluate the impact of YP[4] is on program expenditure. The DEEWR data (4) consists of total annual program expenditure for each individual (Job Seeker Account expenditure, Work for the Dole expenditure and wage subsidies) as well as information on Job Network job placement outcomes.

[4] Most of the interviews were conducted in a two-month window around the one- and two-year anniversaries of entry to the trial—for Year 1, 96.9 per cent of the interviews were conducted in a window 11 to 13 months after entry to the trial; and for Year 2, 93.3 per cent were interviewed at 23 to 25 months after entry to the trial.

Quarterly returns (5) were provided by YP[4] case managers for treatment group members only. We make use of information collected in the quarterly returns where we investigate heterogeneity in impact of YP[4] by treatment dosage. To do this, we construct a measure of treatment duration based on the length of time the case manager maintained contact with the participant.

3.2 Outcomes examined

For the analysis of program impacts using only Centrelink administrative data, we examine three outcomes over Years 1 to 3:

1. Time spent on income support payments in the 12-month period (Centrelink administrative data)
2. Total value of benefit payments in the 12-month period (Centrelink administrative data)
3. Employment assistance program expenditure in the 12-month period. Program expenditure comprises Personal Support Program (PSP), Job Placement, Education and Training (JPET), Job Services Australia (JSA) expenditure, and other program expenditure. (DEEWR data; available only in Years 1 and 2).

For the analysis that incorporates the YP[4] interview data, we are able to examine an expanded set of outcomes. This set comprises the three outcomes described above, plus thirteen extra measures:[5]

Employment

4. Whether employed at anniversary of entry into the YP[4] trial (based on Centrelink administrative data for those on income support at the anniversary date, and based on YP[4] interview data for other participants).

Housing

5. Whether have housing at anniversary of entry into the YP[4] trial (based on YP[4] interview data).
6. Whether ever 'slept rough' in the past 12 months (based on YP[4] interview data).
7. Whether housing is suitable at anniversary of entry into the YP[4] trial (based on YP[4] interview data).
8. Number of moves in the past 12 months (based on YP[4] interview data).

Health and wellbeing

9a/b. Rating of health at anniversary of entry to the YP4 trial (based on YP4 interview data).

[5] Information on the interview questions from which these variables were derived is provided in Appendix Table A4.

10a/b. Rating of well-being at anniversary of entry to the YP4 trial (based on YP4 interview data).

Community engagement

11. Whether has a good connection with at least one family member, someone can talk to and/or someone could ask for practical assistance (based on YP[4] interview data; available only in Year 2).
12. Whether participates in a community activity (based on YP[4] interview data).
13. Number of community services used in the past year (based on YP[4] interview data).
14. Whether had difficulty accessing a community service in the past three months (based on YP[4] interview data).

These outcomes are of course not exhaustive of those relevant to full assessment of the impacts of the YP[4] program. More comprehensive data on labour market outcomes, including hours of work, wage rates, both at anniversary of entry into the program and throughout the intervening year, would be particularly valuable. Nonetheless, a reasonably large number of outcomes are able to be assessed using the available data, allowing reasonably strong inferences to be made about the effects of the YP[4] model of coordinated service delivery.

3.3 Sample

The final sample used in estimation is arrived at after restricting to a balanced panel and taking account of non-response. The restriction to balanced panels is necessary because we use a matching method as the main approach for estimating the impact of the YP[4] program. Having a balanced panel reduces the sample size where Centrelink administrative data are used because data are missing for participants who move outside Victoria during the trial. For the analysis using the YP[4] interview data, the sample size for a balanced panel is further reduced by interview non-response. Since the interview response rate in Year 3 was very low, for analysis using the YP[4] interview data our balanced panel is derived just from participants who responded to both the Year 1 and Year 2 interviews—hence the restriction to Year 1 and Year 2 outcomes when using the interview data.

Table 2 presents the number of sample members in the treatment group (J) (including those allocated reserved places who are defined as being in the reserved treatment group (R)), and in the control group (S). For the analysis of the impact of YP[4] using the Centrelink data the sample is 355 persons; and for the analysis using the YP[4] interview data we have 198 persons. The restriction to a balanced panel can be seen to result in a loss of about 20 per cent of observations when the Centrelink data are used, and about 50 per cent of observations when the YP[4] interview data are also used. Given the small sample size, we evaluate impacts only for all four sites combined. Although

there may be variation in impact of the YP[4] program across the sites, and there is possibly interest in knowing about these differences, sample sizes are not sufficient to allow reliable inferences to be made for individual sites.[6]

Table 2: Sample sizes in the balanced panels

	Treatment group			Control group (S)
	Reserved (R)	Non-reserved	Total (J)	
Panel A: Centrelink data balanced panel				
Frankston	10	33	43	38
Cheltenham	8	34	42	45
Bendigo	8	47	55	48
Central Melbourne	13	36	49	35
Total—All sites	39	150	189	166
Percentage of total sample	78.0	80.6	80.1	87.8
Panel B: Interview data balanced panel				
Frankston	5	19	24	18
Cheltenham	5	20	25	29
Bendigo	7	33	40	33
Central Melbourne	8	14	22	17
Total—All sites	25	86	111	97
Percentage of total sample	50.0	46.2	47.0	51.3

4. Empirical analysis

This section presents the results from empirical analysis of the impact of the YP[4] trial. We begin in Section 4.1 by presenting estimates assuming that randomisation between treatment and control groups was achieved. In Section 4.2 it is shown, however, that the treatment and control groups cannot be treated as randomly assigned. Hence we move on to apply a quasi-experimental matching method to derive estimates of the impact of the YP[4] trial. Section 4.3 explains the methodology, and Section 4.4 presents results. Finally, in Section 4.5, as an alternative method of evaluating the impact of receiving case management services, we present findings from analysis of the effect of duration of participation in YP[4] on outcomes for treatment group members.

The impact of YP[4] that is identified in this study is the *average effect of intention to treat on the treatment group*. First, the scope of the YP[4] trial is necessarily restricted to the population group chosen as eligible to participate. Hence, empirical analysis of the effects of the trial will identify the effect of receiving treatment (case management) on the treated (a population with the same

[6] Appendix Table A1 provides details on the distribution of the entire sample of participants who entered the trial, disaggregated by treatment assignment, site and quarter of entry into the trial. Appendix Table A2 compares the demographic characteristics of the balanced panel samples with the full sample. However, only a limited number of characteristics are available for those not in the Centrelink data balanced panel or interview data balanced panel. Tests of statistical significance of differences in characteristics (not reported) indicate that those not in the balanced panels are more likely to be male and to have entered the trial prior to the fourth quarter of 2005.

characteristics as participants in the YP[4] trial). Second, duration of participation in the YP[4] trial differs between members of the treatment group. Hence, it is the effect of the *intention* to provide the treatment (case management), rather than the actual provision of the full treatment to all members of the treatment group, that can be identified.

4.1 Impact estimates assuming random assignment

Impact estimates of the YP[4] trial obtained by comparing mean outcomes for treatment group members with mean outcomes for control group members are presented in Table 3. These are valid estimates of the effect of case management in YP[4] where assignment between treatment and control groups is random. 'Difference' is the difference between outcomes for the treatment and control groups. The treatment group observations used in Table 3 exclude the reserved group, since it is known that this group were not assigned randomly. Panel A shows outcomes derived from Centrelink administrative data, and Panel B shows the outcomes also using YP[4] interview data.

The message from Table 3 is straightforward. We find no evidence of significant effects of joined-up case management on outcomes. For no outcome variable is there a difference that is significant at the five per cent level.

Table 3: Mean values of outcomes variables and raw differences between treatment and control groups—Balanced panel samples

	Mean of treatment group (J)	Mean of control group (S)	Difference (J – S)	t-statistic
Panel A: Centrelink administrative data				
Year 1				
1. Numbers of days on income support	323.2	315.2	8.1	0.88
2. Total Centrelink payments ($)	7,717.1	7,552.6	164.5	0.53
3. DEEWR program expenditure ($)	1,413.3	1,262.6	150.7	0.77
Year 2				
1. Numbers of days on income support	268.5	273.0	-4.5	-0.33
2. Total Centrelink payments ($)	7,389.7	7,636.2	-246.5	-0.50
3. DEEWR program expenditure ($)	1,380.9	1,186.8	194.1	1.01
Year 3				
1. Numbers of days on income support	250.4	232.4	18.0	1.17
2. Total Centrelink payments ($)	7,522.8	7,255.7	267.2	0.46
Panel B: Centrelink administrative data and YP[4] interview data				
Year 1				
1. Numbers of days on income support	337.5	325.2	84.5	1.13
2. Total Centrelink payments ($)	7,814.8	8,029.9	3,203.9	0.53
3. DEEWR program expenditure ($)	1,402.5	1,297.2	1,790.4	0.41
4. Employed at anniversary of entry to trial	0.10	0.11	0.31	0.06
5. Housed at anniversary of entry to trial	0.36	0.51	0.50	2.04
6. Ever 'slept rough' in past 12 months	0.37	0.27	0.45	1.55
7. Housing is suitable	0.69	0.66	0.48	0.44
8. Number of moves in past 12 months	5.07	2.97	3.73	1.85
9a. Self-rated health good	0.50	0.48	0.50	0.16
9b. Self-rated health bad	0.19	0.15	0.36	0.75
10a. Self-rated wellbeing good	0.36	0.45	0.50	1.19
10b. Self-rated wellbeing bad	0.23	0.20	0.41	0.41
12. Participates in community activities	0.25	0.22	0.41	0.54
13. Number of services used in past 12 months	3.66	4.00	-0.34	1.17
14. Difficulty accessing services	0.28	0.33	-0.05	0.71
Year 2				
1. Numbers of days on income support	285.8	304.9	109.7	1.19
2. Total Centrelink payments ($)	7711.9	8946.2	4840.5	1.87
3. DEEWR program expenditure ($)	1475.3	1097.3	1625.0	1.43
4. Employed at anniversary of entry to trial	0.15	0.12	0.32	0.70
5. Housed at anniversary of entry to trial	0.42	0.52	0.50	1.37
6. Ever 'slept rough' in past 12 months	0.26	0.19	0.40	1.07
7. Housing is suitable	0.68	0.73	0.45	0.84
8. Number of moves in past 12 months	3.46	1.90	2.54	1.44
9a. Self-rated health good	0.44	0.53	0.50	1.25
9b. Self-rated health bad	0.18	0.14	0.35	0.79
10a. Self-rated wellbeing good	0.44	0.57	0.50	1.86
10b. Self-rated wellbeing bad	0.17	0.20	0.40	0.44
11. Has support from family or communities	0.92	0.95	0.23	0.64
12. Participates in community activities	0.25	0.34	0.48	1.40
13. Number of services used in past 12 months	3.84	4.17	-0.34	1.13
14. Difficulty accessing services	0.29	0.32	-0.03	0.41

Note: See Appendix Table A4 for definitions of the outcome variables.

4.2 Tests of randomisation

As described in Section 2, randomisation in the YP4 trial is likely to have been compromised. This implies that differences in mean outcomes of the treatment and control groups are not valid estimates of the impact of YP4. We therefore undertake tests of randomisation by comparing initial (entry date) characteristics of the treatment and control groups. Individuals who were assigned

reserved treatment group places (R) are excluded from these tests, but all other participants for whom recruitment and referral data are available are included. Randomisation is tested for each demographic characteristic by calculating the difference in its mean between treatment and control groups. Results of these tests using the YP[4] recruitment and referral data (1) are reported in Table 4.

Table 4: Mean values of demographic characteristics and tests of random assignment

	Mean of treatment group (J)	Mean of control group (S)	Difference (J-S)	t-statistic
Age (years)	23.24	22.92	0.32	0.70
Male	0.72	0.58	0.14	2.82
Education: Below Year 10	0.28	0.23	0.06	1.27
Education: Year 10	0.51	0.43	0.08	1.60
Education: Above Year 10	0.17	0.31	-0.14	-3.20
Education: Not known	0.03	0.03	0.00	0.03
Immigrant	0.06	0.11	-0.05	-1.81
Immigrant status unknown	0.04	0.04	0.01	0.29
Indigenous	0.09	0.02	0.06	2.80
Ex-offender (previously served time in jail)	0.44	0.24	0.20	4.12
Ex-offender status not known	0.03	0.03	0.01	0.33
Disability in past 12 months	0.45	0.45	0.00	0.04
Homeless at entry into trial	0.42	0.41	0.01	0.13
Homeless status at entry into trial unknown	0.04	0.04	0.01	0.29
DEEWR program expenditure in past 12 months ($)	1,260.97	1,372.46	-111.49	-0.63
DEEWR program expenditure missing	0.06	0.03	0.03	1.48
Number of days on income support in past 12 months	244.85	253.68	-8.83	-0.60
Number of days housed in past 12 months[a]	130.27	158.47	-28.20	-1.81
Number of moves in past 12 months	2.65	2.32	0.33	1.52
Numbers of moves unknown	0.08	0.04	0.04	1.61
Entered YP[4] trial in Quarter 1 2005	0.05	0.02	0.03	1.78
Entered YP[4] trial in Quarter 2 2005	0.52	0.3	0.23	4.54
Entered YP[4] trial in Quarter 3 2005	0.18	0.15	0.03	0.76
Entered YP[4] trial in Quarter 4 2005 or in 2006[b]	0.25	0.53	-0.29	-5.95

Notes: Sample size is 375. [a] 'Housed' if Centrelink records indicate the participant was in private rental housing or public housing in Centrelink record. [b] Only two people entered the trial in 2006, and both entered in January.

For six variables there is a significant difference between mean levels for the treatment and control groups: gender, indigenous status, whether highest level of education attainment is above year 10, whether ex-offender, and whether entered the trial in Quarter 2 2005 or Quarter 4 2005 or later. A Hotelling T test of joint significance rejects at the one per cent level of significance the hypothesis that treatment and control groups have the same mean levels of characteristics. Hence, we infer that randomisation was not successfully implemented.

4.3 Matching method

To address non-random assignment of trial participants, we employ a quasi-experimental matching method (see Borland et al., 2005). Fundamentally, this involves comparing outcomes for the treatment group who received case management and a matched control group who did not receive case management.

For the quasi-experimental matching method to be a valid estimator of the treatment effect, it is sufficient that the following two assumptions hold (Rubin, 1979):

(a) Conditional Independence Assumption (CIA) - Conditional on a set of observable variables (X), participation in treatment is unrelated to outcomes in the absence of treatment; and

(b) Common support assumption - For each possible combination of observable variables there is a non-zero probability of non-participation.

Part (a) effectively requires that matching between treatment and control group observations should be conditional on all variables that affect both participation in case management and outcomes in the absence of case management (Augurzky and Schmidt, 2001). Or, alternatively, after conditioning on the set of X variables, assignment between the treatment and control groups is random. Part (b) is necessary to ensure that, for any treatment group observation, there will be a control group observation with the combination of observable characteristics to which the treatment observation can be matched.

Almost certainly the most important issue in undertaking a matching analysis is to justify why the CIA is likely to hold. Our justification for the CIA is that treatment and control group observations can be matched using a relatively rich set of covariates. Most significantly, it is possible to match on the basis of local labour market and income support payment history. These two factors have been identified as of particular importance in evaluations of matching estimators (for example, Card and Sullivan, 1988, Heckman et al., 1999, and Kluve et al., 2001).[7]

To implement the matching method we use a Propensity Score Model (PSM) approach. Essentially this involves matching treatment and control group observations on the basis of their predicted probability of membership of the treatment group in the YP[4] trial (Rosenbaum and Rubin, 1983).

Stage one of the PSM approach is to estimate a probit model for whether a participant in the YP[4] trial is in the treatment group. Covariates included in the model are: gender; age category; educational attainment; whether an immigrant; Indigenous status; whether an ex-offender (previously served time in jail); whether homeless at time of entry to trial; whether on income support payments in the six months prior to entry to the trial; total time on income support payments in the past 12 months; DEEWR program expenditure in the past 12 months; number of moves in the past 12 months; number of days housed in the past 12 months; whether recorded by Centrelink as having a disability in the past 12 months; trial site at which located; and quarter commenced participation in the trial.

[7] Recent studies for Australia, using other data sources, also establish the importance of labour force history in explaining labour market status. Le and Miller (2001) and Knights et al. (2002) have shown that once labour market history is controlled for, other standard covariates have very little explanatory power for whether a labour force participant is unemployed or employed.

A balancing test is used to find an appropriate functional form of the probit model for assignment to treatment (see Dehejia and Wahba, 1999, 2002, and Smith and Todd, 2005). Specifically, a post-matching balancing test was used to derive the model (see Table A3 in the Appendix). This balancing test evaluates whether there is a significant difference between the mean value of a characteristic for treatment group members and the weighted average of that characteristic for control group members (with weights corresponding to the number of times a control observation is used in matching). For the analysis using Centrelink administrative data, balance was achieved for 26 out of 28 matching variables, but could not be achieved for one of the dummy variables for the number of moves (moved 2-3 times in the past year)[8] and one of the location dummy variables (Cheltenham). This is likely to be due to randomization having been compromised, in particular at the Cheltenham site, where 30 out of 31 participants who entered in quarter 2, 2005 were assigned to group J and all participants who entered in quarter 4, 2005 were assigned to group S. For the analysis also using the YP[4] interview data, balance was achieved for all 28 matching variables.

Stage two of the PSM is to match treatment and control group observations. A formal description of the estimated effect of YP[4] using the 'basic' matching estimation method is:

$$\tau = \sum_{i \in T} \left[Y_{Ti} - \sum_{j \in C} w(i,j) Y_{Cj} \right] \qquad (1)$$

where Y is the outcome variable, T and C denote treatment and control groups respectively, and $w(i,j)$ are the weights placed on the j^{th} potential control group observation in constructing a comparison for the i^{th} treatment group observation.

Matching of treatment and control observations involves three steps. First, each treatment observation is matched to a weighted average of control observations in a ninety-five per cent confidence interval (caliper method). Second, a difference in outcome is calculated between each treatment observation and its matched weighted average of control observations. Third, the aggregate effect of case management in YP[4] is calculated as the average of the difference in outcomes across all treatment observations. Matching is undertaken using the linear predicted score from PSM, which is preferred to the predicted probability as this allows symmetry in selection of control observations using the caliper method. Re-sampling of control observations across different treatment observations is allowed. Weights for control observations are derived using kernel weighting:

[8] However, the difference in the average numbers of moves between the treatment and matched control groups is not statistically significant at the ten per cent level.

$$w(i,j) = \frac{G^{ij}}{\sum_{j \in C} G^{ij}}$$
$$\text{and } G^{ij} = G\left(\frac{X_i\hat{\beta} - X_j\hat{\beta}}{a_{95\%}}\right) \quad (2)$$

where G^{ij} is the kernel for i^{th} treatment and j^{th} control observations for the male sample, $X_i\hat{\beta}$ and $X_j\hat{\beta}$ are linear predicted scores for the respective treatment and control observations, and $a_{95\%}$ represents the use of a 95 per cent confidence interval bandwidth around $X_i\hat{\beta}$. In this approach the bi-weight kernel is used.

Validity of the matching estimator requires that the CIA and common support assumptions should hold. There is no formal test for the CIA; instead above we have provided justification for why we believe the assumption is satisfied. The common support assumption however can be assessed empirically. Using the basic matching method, only eight out of 355 treatment observations cannot be matched to a control group observation. The average number of control observations matched per treatment observation is 73 for the Centrelink administrative data and 42.5 for the YP[4] interview data. The mean number of times each control observation was used in matching is 1.14 for the Centrelink administrative data and 1.06 for the YP[4] interview data. Only one control observation was not used in matching

4.4 Matching method results

Estimates of the impact of the YP[4] program derived using the matching method are presented in Table 5. Here, reserve place treatment group members are included in the treatment group sample because selection is explicitly acknowledged to be non-random and the matching process is designed to account for this non-randomness.[9]

Once again, there is an absence of evidence of significant effects of joined-up service delivery on outcomes. No significant effect is found for any outcome using the administrative data (Panel A). When we examine the sample using interview data (Panel B), there are two outcomes where program participation has a significant effect at the five per cent level. Receiving case management is shown to have reduced the number of community services used in the first year after entry to the trial, which could perhaps be interpreted as indicative of better matching of services to individual need. Receiving case management is also shown to reduce the amount of income support payments received by $1,785 over the second year after entry to the trial (or about $70 per fortnight).

[9] Similar results are obtained excluding the reserve place treatment group members.

However, it is important to recall that the balanced panel using interview data represents only about 50 per cent of the population of trial participants; and when we examine income support payments received for the larger group of trial participants using the Centrelink balanced panel data, there is no significant effect of case management. For no other outcome (for either balanced panel) is there evidence of a significant effect of receiving case management in the YP[4] trial.

Table 5: Matching method impact estimates—Balanced panel

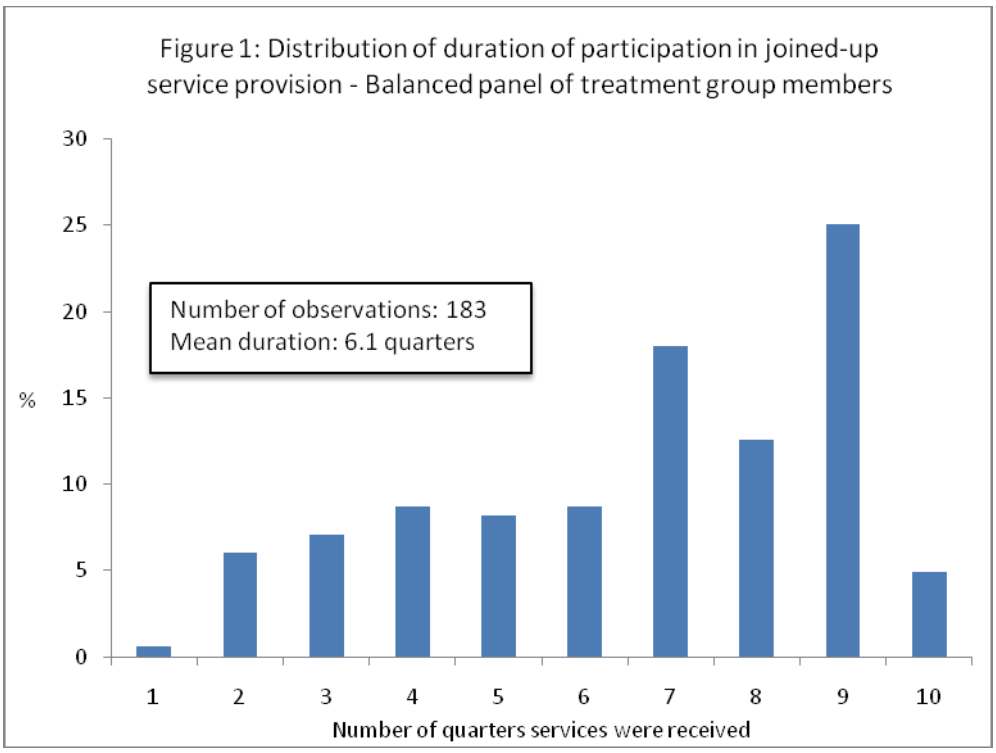
	Year 1		Year 2		Year 3	
	Difference (J-S)	Standard error	Difference (J-S)	Standard error	Difference (J-S)	Standard error
Panel A: Centrelink administrative data						
1. Numbers of days on income support	10.51	13.64	-2.64	17.96	10.74	19.74
2. Total Centrelink payments ($)	258.97	452.89	91.23	602.16	669.95	653.99
3. DEEWR program expenditure ($)	353.2	192.82	-39.92	262.2		
Panel B: Centrelink administrative data and YP[4] interview data						
1. Numbers of days on income support	-3.44	10.83	-37.94	19.62		
2. Total Centrelink payments ($)	-749.79	627.11	-1,785.05	881.5		
3. DEEWR program expenditure ($)	273.89	271.54	321.2	406.2		
4. Employed at anniversary of entry to trial	0.00	0.07	0.01	0.08		
5. Housed at anniversary of entry to trial	-0.10	0.12	0.01	0.11		
6. Ever 'slept rough' in past 12 months	-0.05	0.11	0.06	0.09		
7. Housing is suitable	0.12	0.12	0.05	0.11		
8. Numbers of moves in past 12 months	1.52	1.27	1.58	1.21		
9a. Self-rated health good	-0.09	0.11	-0.05	0.12		
9b. Self-rated health bad	0.06	0.07	0.06	0.07		
10a. Self-rated wellbeing good	-0.05	0.11	-0.07	0.12		
10b. Self-rated wellbeing bad	-0.03	0.1	-0.12	0.11		
11. Has support from family or communities	NA	NA	0.09	0.08		
12. Participates in community activities	0.07	0.09	-0.05	0.11		
13. No. of services used in past 12 months	-0.81	0.41	-0.27	0.45		
14. Difficulty accessing services	-0.22	0.11	-0.14	0.11		

Note: See Appendix Table A4 for definitions of the outcome variables.

4.5 Impact estimates by duration of participation in case management

There was considerable variation in the duration that treatment group members spent receiving case management in the YP[4] trial. This can be seen from Figure 1 which shows that while over 35 per cent received case management for two years or more, there were also about 30 per cent who did not receive the intended minimum of eighteen months.

One would expect that treatment group members who ceased case management early, or never even commenced case management, would have had lower potential to benefit from case management. Similarly, those who had only infrequent contact with their case manager might be expected to benefit less from case management.



Although our primary objective has been to examine the overall effect of the offer of case management on the treatment group, we can also explore whether outcomes differ by duration of participation in case management. This can be done by estimating a regression model that includes dummy variables for duration of participation in case management:

$$E\left[Outcome\right] = \mathbf{X}\beta + \sum_{i=1}^{I} D_i \gamma_i \qquad (3)$$

We use the total time spent on income support payments in the third year after entry to YP[4] as the outcome measure. Estimation is on members of the treatment group only, since the focus is on whether time on income support differs by level of treatment 'dosage'. The coefficients on the duration dummies D_i provide estimates of the different impact of treatment on the outcome depending on the duration of participation in case management. We estimate specifications with six categories of duration of treatment in YP[4]: up to one day (the reference category); more than one day and less than 1 year; 1 to 1.5 years; 1.5 to 2 years; and more than 2 years. We restrict outcomes to the Centrelink administrative data because it is available for all three years for most participants, and because differences in impact by duration of participation can only satisfactorily be conducted when the longer-term (third-year) outcomes are examined.

It is important to note that our estimates of the effect of duration of participation in YP[4] will only be valid under the assumption that treatment group members were randomly assigned to different

lengths of case management. However, it seems likely that length of case management is correlated with characteristics that may also be determinants of the outcomes, such as a participant's level of motivation. Where more motivated treatment group members take advantage of case management for longer periods of time, the effect of duration of case management on outcomes will confound both the effects of motivation and duration of treatment.

Table 6: Effect of YP[4] services on total time spent on income support payments in the third year after commencement of program participation, by length of time received services—OLS regression model estimates

	Coefficient estimate	Standard error	t-statistic
Length of time received YP[4] services (*Reference category*: Less than 1 day)			
More than 1 day but less than 1 year	-2.46	34.07	-0.07
At least 1 year but less than 1.5 years	-33.14	35.78	-0.93
At least 1.5 years but less than 2 years	-41.79	33.67	-1.24
2 years or more	-62.07	35.19	-1.76
Male	-38.81	25.50	-1.52
Age group (*Reference category*: 18-20 years)			
21-24 years	28.68	26.77	1.07
25 years and over	28.48	28.09	1.01
Education (*Reference category*: Below Year 10)			
Year 10	-18.32	24.58	-0.75
Above Year 10	-41.18	30.11	-1.37
Immigrant	15.58	41.39	0.38
Indigenous	-21.15	35.59	-0.59
Ex-offender (previously served time in jail)	34.16	22.91	1.49
Homeless at entry to trial	-34.57	21.53	-1.61
Not on income support in the 6 months prior to entry to trial	-26.21	38.28	-0.68
Total time on income support in past 12 months (days)	0.57	0.12	4.88
DEEWR program expenditure in past 12 months (*Reference category*: $0)			
$1 to $1,200	-3.27	26.90	-0.12
$1,201 to $2,000	29.92	31.80	0.94
$2,000 or more	-26.28	28.39	-0.93
Numbers of moves in past 12 months (*Reference category*: 0 or 1)			
2 or 3	10.37	26.84	0.39
4 or 5	4.19	31.21	0.13
6 or more	94.24	39.03	2.41
Numbers of days housed in past 12months (*Reference category*: 0)			
1 to 182 days	69.08	31.23	2.21
183 days to 364 days	-24.01	32.91	-0.73
365 days	-4.28	35.92	-0.12
Disability recorded in past 12 months	-6.29	22.30	-0.28
Site (*Reference category*: Frankston)			
Cheltenham	-4.50	31.97	-0.14
Bendigo	70.09	30.05	2.33
Central Melbourne	-2.63	33.05	-0.08
Date of entry to the YP[4] trial (*Reference category*: Quarter 1 2005)			
Quarter 2 2005	-12.70	37.97	-0.33
Quarter 3 2005	5.12	44.20	0.12
Quarter 4 2005 or later	-7.89	44.65	-0.18
Constant	96.43	63.53	1.52

Notes: Sample comprises the balanced panel for which all three years of Centrelink administrative data is available. Sample size is 187. R-squared is 0.2997.

Table 6 shows the results from estimation of the regression model. The coefficient estimates on the duration variables do show that increasing time in case management is associated with decreased time on income support payments—for example, receiving two years or more of case management is associated with a reduction in time spent on income support payments of about three weeks in the third year after entry into the program. However, none of the estimated effects is statistically significant at the five per cent level. It is also important to take into account that these findings may reflect 'selection' effects.

5. Concluding comments

Our study is notable in several important ways—in being a randomised control trial; in its analysis of an intervention targeted at homeless jobseekers; in considering a new model for case management, YP[4], that sought to provide a more comprehensive approach to coordinating service delivery; and in the breadth of the outcome measures and extended time horizon that we use to examine the impact of YP[4].

The evidence from the YP[4] trial is that the new model for coordination of service delivery did not make a difference to outcomes for young homeless jobseekers. We have considered a large number of outcome measures covering many aspects of well-being; and have analysed those outcomes over an extended period after participants began receiving joined up service delivery. We find no evidence of significant effects of joined up service delivery on outcomes. This applies irrespective of whether random assignment to treatment is assumed or non-random assignment is controlled for using the matching method.

Our findings provide two important contributions to the international literature on program design for jobseekers and the homeless. One lesson is that 'you get what you pay for'. The implementation of case management in the YP[4] trial appears to have made it a relatively minimal intervention. Indeed, this is borne out by the absence of significant differences between the matched treatment and control groups in utilisation of employment services and in difficulty accessing community services. But when dealing with a highly disadvantaged population, minimal interventions are unlikely to effect substantial improvements in outcomes. To give some extra context to this point, we know that an extra year of schooling for an average student in Australia will add 5 to 10 per cent to labour market earnings (see for example, Borland et al., 2000); yet here we are talking about a much more disadvantaged population group than average, and an intervention that is a tiny fraction of what is provided through a year of schooling. That minimal interventions are likely to have minimal effects is a point that has been made by Heckman et al. (1999) in their review of the effects of active labour market programs; and has been exemplified in Australia by the Department of Family and

Community Services' trial to evaluate the effect of extra case management for the very long-term unemployed (see Breunig et al., 2003).

A second lesson is that administration matters. Much about whether a program will be successful depends on its implementation. Partly this is about the program actually being received by its intended recipients. We have seen in the case of the YP[4] trial that 20 per cent of the treatment group never met with their case manager; and over 50 per cent met with a case manger on average only once every 6 months during the trial. Implementation is also about case managers effectively carrying out a program. That this may have been an issue in the YP[4] trial is suggested by our finding that there was no effect of the program utilisation or difficulty of accessing services.

The finding from the YP[4] trial that administrative effects matter for program outcomes again builds on a small existing literature. On the issue of the program being carried out, Borland and Tseng (2010), for example, have documented how labour market programs in Australia that were regarded as being universal, often in fact had far from 100 per cent take-up rates, mainly due to a failure by administrators to require all eligible payment recipients to undertake the programs. On the issue of capacity to carry out a program, an example is the study by Lechner and Smith (2007) which found that case managers in Switzerland appeared to have limited knowledge of which eligible welfare payment recipients would be most likely to benefit from available employment services and labour market programs; and suggested that using statistical methods of assignment achieved superior outcomes.

While this study did not find evidence of significant effects of the YP[4] intervention, the commitment of the welfare and government agencies involved to credible evaluation of the program should be commended. It is one of the very few attempts in this policy arena to conduct a randomised controlled trial and thereby produce rigorous scientific evidence on the impact of the trialled program on outcomes for the target group. It thus represents an important step in social policy development in Australia, providing a number of valuable lessons for future social policy evaluation. The difficulties confronted in this trial should be instructive not only on the potential problems in conducting randomised controlled trials, but indeed also on the absence of meaningful information on efficacy of social policy interventions that are implemented without such trial: as difficult as it is to identify program effects on participants when explicitly attempting to create a 'counterfactual' by having a control group, it will necessarily be even harder in the absence of such efforts.

6. Appendix

Table A1: Sample sizes by treatment assignment, site and quarter of entry into YP[4]

		Quarter of entry					Total
		Quarter 1 2005	Quarter 2 2005	Quarter 3 2005	Quarter 4 2005	Quarter 1 2006	
Site 1	S	0	17	6	18	0	41
Frankston	J	4	21	8	9	0	42
	R	0	7	3	0	0	10
	Total	4	45	17	27	0	93
Site 2	S	0	1	11	35	2	49
Cheltenham	J	2	30	12	0	0	44
	R	2	7	2	1	0	12
	Total	4	38	25	36	2	105
Site 3	S	2	19	8	26	0	55
Bendigo	J	2	23	8	20	0	53
	R	4	5	1	0	0	10
	Total	8	47	17	46	0	118
Site 4	S	1	19	4	20	0	44
Central Melbourne	J	1	23	6	17	0	47
	R	8	6	2	2	0	18
	Total	10	48	12	39	0	109

Notes: *S*–Control group; *J*–Treatment group (unreserved); *R*–Treatment group (reserved).

Table A2: Comparison of means of demographic characteristics of balanced panels and full sample

	Centrelink data balanced panel	Interview data balanced panel	Full sample
Age (years)	23.28	23.24	23.17
Male	0.64	0.59	0.66
Education: Below Year 10	0.27	0.27	0.25
Education: Year 10	0.47	0.43	0.46
Education: Above Year 10	0.26	0.30	0.25
Education: Not known	0.01	0.00	0.03
Immigrant	0.09	0.06	0.09
Immigrant status unknown	0.01	0.00	0.04
Indigenous	0.06	0.06	0.06
Ex-offender (previously served time in jail)	0.35	0.33	0.34
Ex-offender status not known	0.01	0.00	0.03
Entered YP[4] trial in Quarter 1 2005	0.05	0.04	0.06
Entered YP[4] trial in Quarter 2 2005	0.40	0.40	0.42
Entered YP[4] trial in Quarter 3 2005	0.17	0.14	0.17
Entered YP[4] trial in Quarter 4 2005 or in 2006[b]	0.37	0.42	0.35

Table A3: Probit model for matching method propensity score

	Centrelink administrative data		Centrelink administrative data and interview data	
	Coefficient estimate	Standard error	Coefficient estimate	Standard error
Male	0.21	1.13	0.51	1.99
Age group (*Reference category*: 18-20 years)				
21-24 years	0.31	1.41	0.29	0.90
25 years and over	0.17	0.78	0.19	0.61
Education (*Reference category*: Below Year 10)				
Year 10	0.35	1.87	0.30	1.03
Above Year 10	-0.05	-0.24	0.05	0.16
Immigrant	-0.20	-0.72	-0.93	-1.99
Indigenous	1.31	3.18	0.90	1.62
Ex-offender (previously served time in jail)	0.36	2.04	0.65	2.63
Homeless at entry to the YP[4] trial	0.08	0.5	0.00	-0.02
Not on income support in the 6 months prior to entry to the YP[4] trial	-0.63	-1.94	-0.97	-1.81
Total time on income support in past 12 months (*Reference category*: Less than 183 days)				
183 to 364 days	-0.55	-2.02	-0.62	-1.36
365 days	-0.15	-0.58	-0.34	-0.77
DEEWR program expenditure in past 12 months (*Reference category*: $0)				
$1 to $1,200	0.01	0.04	0.05	0.16
$1,201 to $2,000	-0.55	-2.39	-0.67	-2.15
$2,000 or more	-0.09	-0.38	-0.78	-2.20
Numbers of moves in past 12 months recorded by Centrelink (*Reference category*: 0 or 1)				
2 or 3	0.09	0.47	-0.29	-1.03
4 or 5	0.10	0.42	-0.34	-1.01
6 or more	0.62	1.8	0.34	0.67
Numbers of days housed in past 12 months (*Reference category*: 0)				
1 to 182 days	-0.26	-1.16	0.17	0.51
183 days to 364 days	-0.14	-0.68	-0.05	-0.17
365 days	-0.12	-0.54	-0.23	-0.69
Disability recorded in past 12 months	0.06	0.36	-0.19	-0.86
Site (*Reference category*: Frankston)				
Cheltenham	0.03	0.12	0.02	0.06
Bendigo	0.13	0.61	0.14	0.45
Central Melbourne	0.24	0.99	0.29	0.79
Date of entry to the YP[4] trial (*Reference category*: Quarter 1 2005)				
Quarter 2 2005	-1.02	-2.03	-4.65	-0.02
Quarter 3 2005	-1.39	-2.7	-4.77	-0.02
Quarter 4 2005 or later	-2.07	-4.15	-5.70	-0.02
Constant	1.21	2.03	5.28	0.02
Numbers of observations	355		208	
Pseudo R-squared	0.2181		0.2979	

Table A4: Definitions of outcome variables

1. Number of days on income support	Number of days in receipt of an income support payment (welfare benefit) in the 12-month period, as recorded in the Centrelink administrative data.
2. Total Centrelink payments	Total value of government benefits received in the 12-month period, as recorded in the Centrelink administrative data.
3. DEEWR program expenditure	Total expenditure on employment-assistance programs in the 12-month period, as recorded in the DEEWR administrative data.
4. Employed at anniversary of entry to trial	For those on income support payments at the anniversary of entry into the YP[4] trial, recorded in the Centrelink administrative data as reporting earnings in the fortnight of the anniversary. For those off income support, obtained from responses to the question *Are you currently in paid work?* administered in the annual interview.
5. Housed at anniversary of entry to trial	Derived from the annual interview question *Where are you living at the moment?* The variable is equal to one if housed in private rental accommodation or in public housing, and is equal to zero otherwise.
6. Ever 'slept rough' in past 12 months	Derived from the annual interview question *Have you stayed in any of the following types of accommodation in the past 12 months?* The variable is equal to one if the response option 'Sleeping rough (street/squat/carpark)' was selected, and is zero otherwise.
7. Housing is suitable	Derived from the annual interview question *How would you rate the suitability of your present living arrangements?* The variable is equal to one if the response options 'highly suitable' or 'suitable' were selected, and zero if the response option 'unsure', 'unsuitable' or 'extremely unsuitable' were chosen.
8. Number of moves in past 12 months	Obtained from the annual interview question *How many moves have you made in the past 12 months?*
9a. Self-rated health good 9b. Self-rated health bad	Derived from the annual interview question *How would you rate your overall health at the moment?* The variable 'self-rated health good' is equal to one if the response options 'very good' or 'good' were selected, and is equal to zero if the response options 'average', 'not good' or 'poor' were selected. The variable 'self-rated health bad' is equal to one if the response options 'not good' or 'poor' were selected, and is equal to zero otherwise.
10a. Self-rated wellbeing good 10b. Self-rated wellbeing bad	Derived from the annual interview question *How would you rate your wellbeing at the moment? By wellbeing we mean your mental and emotional health.* The variable 'self-rated wellbeing good' is equal to one if the response options 'very good' or 'good' were selected, and is equal to zero if the response options 'average', 'not good' or 'poor' were selected. The variable 'self-rated wellbeing bad' is equal to one if the response options 'not good' or 'poor' were selected, and is equal to zero otherwise.
11. Has support from family or communities	Variable is equal to one if the participant answered 'yes' to any of the following questions administered in the annual interview: *Do you have a good connection with one or more family members?* *If you were worried about something do you have someone outside your family that you could talk to (not a worker in an agency)?* *If you needed some practical assistance, for example lifting something heavy if you were moving house, do you have someone you could ask for help?*
12. Participates in community activities	Variable is equal to one if the participant answered 'yes' to the annual interview question *Do you participate in community activities such as sports, clubs, or organised groups?*
13. Number of services used in past 12 months	Count of number of the following community services used in the past year: (1) Housing service; (2) Generalist counselling; (3) Financial counselling; (4) Lifeline or other telephone service; (5) Neighbourhood house/community centre; (6) Consumer or tenancy service; (7) Personal development supports; (8) General Practitioner; (9) Community health service; (10) Drug treatment services; (11) Mental health services; (12) Public hospital. The variable ranges from 0 to a potential maximum of 12.
14. Difficulty accessing service	Variable is equal to one if the participant answered 'yes' to the annual interview question *Have you had difficulty accessing services in the past three months?*

7. References

Augurzky, B. and C. Schmidt (2001), 'The propensity score: A means to an end', Discussion Paper no. 271, IZA.

Barrett, G. and D. Cobb-Clark (2000), 'The labour market plans of Parenting Payment recipients: Information from a randomised social experiment', Australian Journal of Labour Economics, 4, 19-205.

Borland, J., P. Dawkins, D. Johnson and R. Williams (2000), 'Returns to investment in higher education', Melbourne Economics of Higher Education Research Program report no.1.

Borland, J., Y. Tseng and R. Wilkins (2005), 'Experimental and Quasi-Experimental Methods of Microeconomic Program and Policy Evaluation', pages 83-118 in Quantitative Tools for Microeconomic Policy Analysis (Canberra, Ausinfo).

Borland, J. and Y. Tseng, 2007. Does a minimum job search requirement reduce time on unemployment payments?: Evidence from the Jobseeker Diary in Australia. Industrial and Labor Relations Review 60, 357-78.

Borland, J. and Y. Tseng (2010), Can mandatory labour market programs improve labour market outcomes for young job seekers?: Threat and participation effects from the Mutual Obligation Initiative in Australia', mimeo.

Breunig, R., D. Cobb-Clark, Y. Dunlop and M. Terrill (2003), 'Assisting the long-term unemployed: Results from a randomized trial', Economic Record, 79, 84-102.

Campbell, S., Horn, M. and Nicholson, T. (2003) *A New Approach to Assisting Young Homeless Job Seekers*, Hanover Welfare Services, Melbourne. Downloadable at https://www.hanover.org.au/component/option,com_docman/task,cat_view/gid,33/dir,DESC/order,name/Itemid,99999999/limit,5/limitstart,20/

Card, D., J. Kluve and A. Weber (2009), 'Active labour market policy evaluations: A meta-analysis', IZA, Discussion paper no.4002.

Card, D. and D. Sullivan (1988), 'Measuring the effect of subsidized training programs on movements in and out of employment', Econometrica, 56, 497-530.

Coventry, L. and Y. Pedrotti (2006) 'Recruitment to YP[4]: An Extensive Description,' Centrelink, mimeo.

Coventry, L. and E. Oiza (2007) 'The Collocation of YP[4] and Centrelink in Bendigo: A case study of partnership in action,' Hanover Welfare Services, Melbourne. Downloadable at <http://www.yp4.org.au/UserFiles/2/File/final%20draft%20colllocation%20paper%20-%20amended.pdf>

Dehejia, R. and S. Wahba (1999), 'Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs', Journal of the American Statistical Association, 94, 1053-1062.

Dehejia, R. and S. Wahba (2002), 'Propensity score matching for nonexperimental causal studies', Review of Economics and Statistics, 84, 151-161.

Fine, M., K. Pancharatnam and C. Thomson (2005), 'Coordinated and integrated human service delivery models', University of NSW, Social Policy Research Centre, Report 1/05.

Fitzpatrick-Lewis, D., Ganann, R., Krishnaratne, S., Ciliska, D., Kouyoumdjian, F. and Hwang, S.W. (2011) 'Effectiveness of Interventions to Improve the Health and Housing Status of Homeless People: A Rapid Systematic Review', BMC Public Health 11(638): doi:10.1186/1471-2458-11-638.

Grace, M. and Gill, P. (2008) *Improving outcomes for homeless jobseekers: YP[4] participant outcomes by amount of case management service received*, Victoria University, Melbourne. Downloadable at <http://www.yp4.org.au/index.asp?menuid=160.010&artid=31>

Heckman, J., R. Lalonde and J. Smith (1999), 'The economics and econometrics of active labor market programs', pages 1865-2097 in O. Ashenfelter and D. Card (eds.) Handbook of Labor Economics Volume 3A (Amsterdam, Elsevier).

Horn, Michael (2004) *A new approach to assisting young homeless job seekers: trial proposal*, Hanover Welfare Services, Melbourne. Downloadable at http://www.yp4.org.au/UserFiles/2/File/Proposal%20Document%20Mar%2004.doc

Kluve, J. (2006), 'The effectiveness of European active labor market policy', IZA, Discussion paper no.2018.

Kluve, J., H. Lehmann, and C. Schmidt (2001), 'Disentangling treatment effects of Polish active labour market policies: Evidence from matched samples', mimeo, IZA.

Knights, S., M. Harris and J. Loundes (2002) 'Dynamic relationships in the Australian labour market: Heterogeneity and state dependence', Economic Record, 78, 284-298.

Le, A. and P. Miller (2001), 'Is a risk index approach to unemployment possible?', Economic Record, 77, 51-70.

Lechner, M. and J. Smith (2007), 'What is the value added by caseworkers?', Labour Economics, 14, 135-51.

Lind, B., D. Weatherburn, S. Chen, M. Shanahan, E, Lanscar, M. Haas and R. De Abreu Lourenco (2002), New South Wales Drug Court Evaluation: Cost-Effectiveness (Sydney, NSW Bureau of Crime Statistics and Research).

Parkinson, S. and Horn, M. (2002) *Homelessness and Employment Assistance: A Research Report Examining the Effectiveness of Assessment and Job Referral Procedures for Income Support Recipients Experiencing Homelessness*, Hanover Welfare Services, Melbourne. Downloadable at <https://www.hanover.org.au/component/option,com_docman/task,cat_view/gid,33/dir,DESC/order,name/Itemid,99999999/limit,5/limitstart,10/>

Perri, S. (2004) 'Joined up government in the Western world in comparative perspective: A preliminary literature review and exploration', *Journal of Public Administration Research and Theory*, 14, 103-138.

Rosenbaum, P. and D. Rubin (1983), 'The central role of the propensity score in observational studies for causal effects', Biometrika, 70, 41-55.

Rosenheck, R.A., Lam, J., Morrissey, J.P., Calloway, M.O., Stolar, M. and Randolph, F. (2002) 'Service Systems Integration and Outcomes for Mentally Ill Homeless Persons in the ACCESS Program', Psychiatric Services, 8(53), 958-66.

Rubin, D. (1979), 'Using multivariate matched sampling and regression adjustment to control bias in observational studies', Journal of the American Statistical Association, 7, 34-58.

Smith, J and P. Todd (2005), 'Does matching overcome Lalonde's critique of nonexperimental estimators?', Journal of Econometrics, 125, 305-53.

Wolfers, J. (2006), 'Did unilateral divorce laws raise divorce rates? A reconciliation and new results?', American Economic Review, 96, 1802-20.