IN CONFIDENCE

VETERANS' HOME CARE PROGRAMME: STAGE TWO – ANALYSIS OF IMPACT ON OTHER PROGRAMMES OF THE DEPARTMENT OF VETERANS' AFFAIRS

BY

ACCESS ECONOMICS

FOR THE

HOUSING & AGED CARE BRANCH

DEPARTMENT OF VETERANS' AFFAIRS

JANUARY 2003

CONTENTS

1. EXECUTIVE SUMMARY

DVA previously engaged Access Economics in August 2002 to assess the impact of VHC on the costs and usage of other DVA programmes. That analysis was conducted to tight deadlines and with a limited run of data (VHC only commenced in early 2001).

The conclusion was that spending was up, but the results were heavily qualified – it was too early to tell whether the higher spending represented the relative 'frailty' of those in the VHC, and/or an initial burst of preventative spending, and/or over-servicing.

An additional 7 months of data has now become available. The analysis has also been enriched by using the Disability Pension (DP) as a proxy for frailty. In addition, the pre-VHC experience has been extended to six months, whereas the previous analysis was limited to three months. That has smoothed out some of the fluctuations in the data and reduced the need for any trimming of outliers.

1.1 OVERVIEW OF RESULTS

The results of the current analysis are as given in Table 1 and Table 2

Table 1 – Cross-tabulation of average spending changes

	Female	Male	Total
Non-transitional	+ $ 33.70	- $ 10.52	+ $ 9.50
Ex-HACC	- $ 109.68	- $ 21.49	- $ 70.27
Total	- $ 28.64	- $ 14.25	- $ 21.32

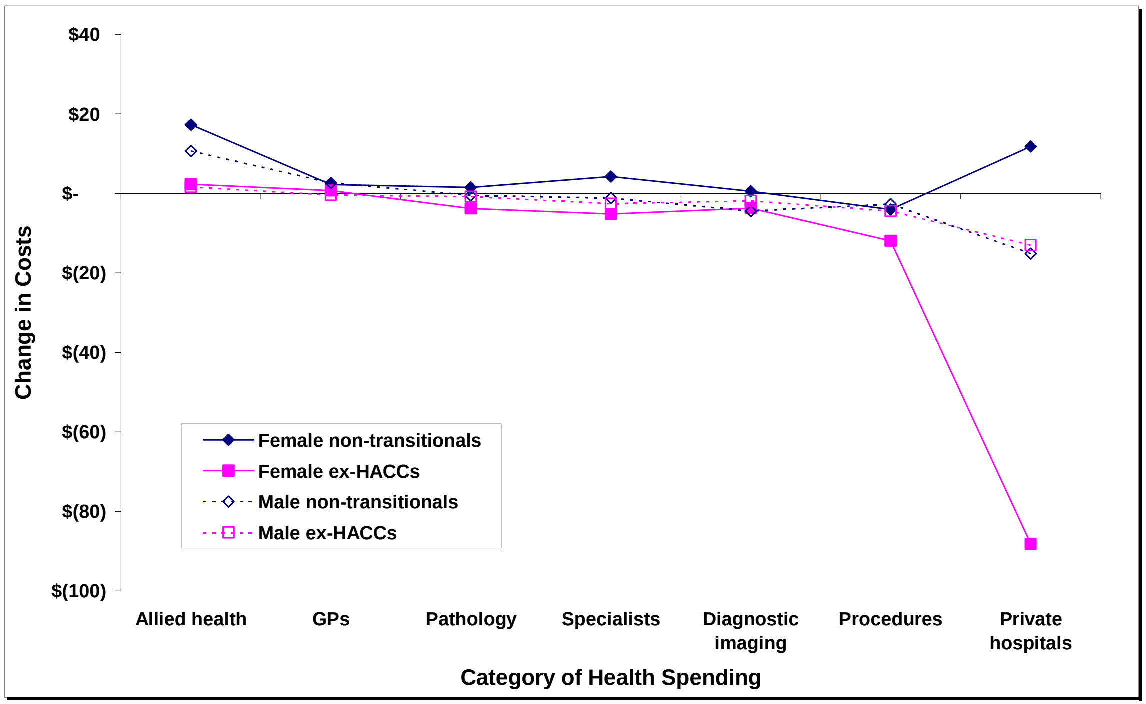
Table 1 shows greater savings for women, and greater savings for ex-HACCs. The latter is somewhat counterintuitive – although VHC doubtless tries to do better than HACC, in many ways it contains elements of a HACC-like programme.

Table 2 – Detail by gender and transitional status

Gender	Non-transitional or ex-HACC	Number	Allied health	GPs	Pathology	Specialists	Diagnostic imaging	Procedures	Private hospitals	Total
Fem.	Non-T	4304	+17.33	+2.23	+1.54	+4.28	+0.56	-4.03	+11.79	+33.70
Fem.	Ex-H	3311	+2.29	+0.74	-3.76	-5.13	-3.68	-11.93	-88.22	-109.68
Male	Non-T	5203	+10.69	+2.67	-0.45	-1.15	-4.44	-2.68	-15.16	-10.52
Male	Ex-H	2675	+1.63	-0.40	-0.81	-2.58	-1.88	-4.43	-13.00	-21.49
Overall		15493	+9.17	+1.61	-0.67	-0.74	-2.44	-5.33	-22.92	-21.32

Table 2 shows that the results here are dominated by the volatility of private hospital spending. Private hospitals are the greatest contributor to the bottom line, as well as generating the largest volatility between classes. Allied health and GP spending go against the trend, registering an increase, especially for non-transitionals.

Chart 1 – Change in Costs by Category



As seen in both tables, our best assessment is that, for equivalent individuals holding gold cards, spending is some 3.6% lower ($21.32 per person per month) in VHC than outside it.

That said, there would be risks in placing too much weight on these results. Indeed, as spending on private hospitals and procedures is also the most volatile element of spending, which weakens the reliance that can be placed on these results. As noted, they remain dominated by the volatility of private hospital spending. There is therefore a reasonable chance that both important results – the overall saving and the emerging evidence of successful 'preventative' spending – may simply be statistical noise.

It is worth noting that the most robust result is an increase in spending on allied health, rather than the overall fall in spending or the large fall in private hospital spending that dominates it.

That said, the overall results point to a saving to the Commonwealth Government from the VHC programme, and one that is growing over time.

1.2 Frailty

The $21.32 saving is made up of two steps. The first two-thirds of the savings comes from the basic model, and the rest comes after extending the basic model to allow for a proxy for frailty – that is, it allows for the possibility that those selected to enter VHC are 'more frail' than those outside VHC, meaning that individuals in VHC are somehow more frail or sick than those outside of VHC.

Using disability pension rates as a proxy for frailty proved significant – allowing for the extra information in the disability variable improves the ability to predict whether someone went into VHC or not.

Chart 2 shows that allowing for frailty has an impact exactly where it would be expected to – on men rather than women (as by and large it is men who receive Disability Pensions), and on non-transitionals rather than ex-HACCs (HACC is a similar program to VHC, so those who are 'more frail' may well have already gone into HACC prior to their transfer into VHC).

Note that the Disability Pension therefore seems to work well as a proxy for 'frailty' for men. The lack of an equivalent proxy for women suggests that these results may understate the true savings under VHC.

Chart 2 – Effect of allowing for frailty by class of person

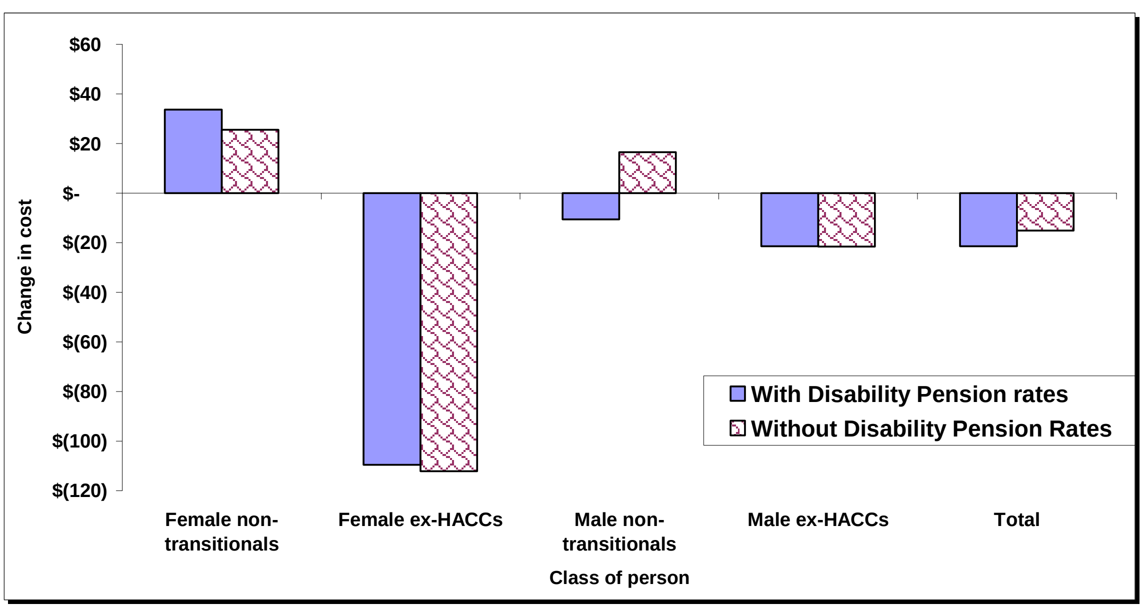
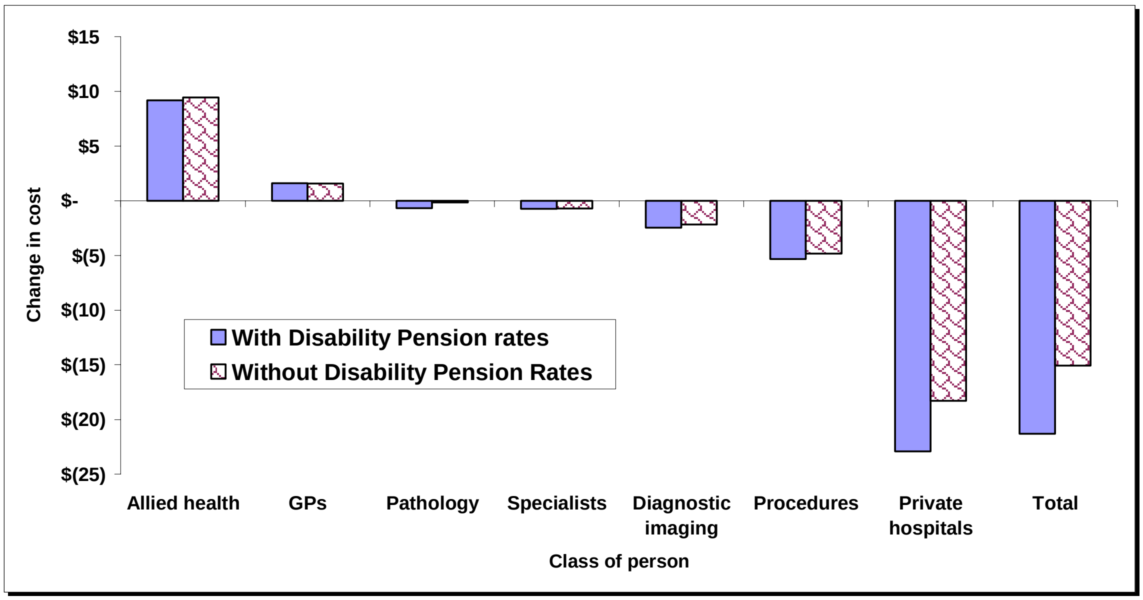


Chart 3 does not necessarily strengthen or weaken the case that using the Disability Pension as a proxy for frailty was correct. It does, however, make the basic point that it is spending on private hospitals which is most volatile.

Chart 3 – Effect of allowing for frailty by category of spending



1.3 PREVENTION RATHER THAN CURE?

The allowance for frailty saw a trend emerge in the results in this report – one consistent with the hypothesis that VHC is seeing more 'preventative' spending early, and that extra investment is paying off via less 'treatment' or 'cure' costs later. Earlier entrants exhibit overall cost savings, and later entrants exhibit cost increases. Four possible explanations of this are:

1. Earlier cohorts (those people who joined the VHC programme earlier) were different in some significant way, such as being more or less frail, or having been 'cherry-picked'. Most of the HACC transferees were moved in June and July of 2001, and in the months leading up to it, and that may affect the results.

2. We have more data for earlier cohorts – and there is a short term cost increase but a long term cost saving (which is only apparent for earlier cohorts).

3. We have more data for earlier cohorts – but our numbers for later cohorts have less confidence associated with them (that is, larger standard errors). For example, individual monthly spending on private hospitals can range up to

$110,000. At the time of our analysis, we had only two months of reliable spending data for 2002-March entrants, yet earlier cohorts had up to 14 months of data.

4. It is the result of chance, with the differences across cohorts essentially driven by the volatile private hospital category.

1.4 The caveats

The small print this time is much the same as last time. There are caveats on the results because:

- *Potentially important information may not be in the available data.* While the data covers a longer period, it contains few extra variables. Unobserved effects such as 'frailty' and 'illness' remain unobserved. The report has described how the methodology allows for some unobserved factors, but may be biased if others are present. This report addresses this using a proxy for 'frailty' in the analysis, but there is no guarantee that the results are still unaffected by other unobserved variables.
- *It is 'one-size-fits-all' modelling.* A total of 20 different models are estimated, covering gender, transitional status (ex-HACC or not), and time. The same set of variables is used in each. No attempt is made to delete insignificant variables.
- *Error estimates are probably too small.* The estimated standard errors do not take into account all potential sources of uncertainty, and so are probably too small. That suggests that the results are probably less definitive than they may otherwise seem.
- *The results are sensitive to the treatment of outliers.* Outliers are handled by using a six month base period (thereby smoothing some volatility in the estimates). However, that does not remove the underlying difficulty that outliers still generate much of the volatility in spending, and so any method of dealing with them runs the risk of throwing away useful information.

1.5 Approach taken

VHC only dates from January 2001. As such the run of data available is relatively limited. There is however unit record data that gives the health spending experience of all veterans and war widows holding gold or white cards, as well as the VHC experience of all individuals in VHC. In addition there is an identifier in the data that allows identification of who is in receipt of the DP and their assessed level of disability.

Given that those in VHC are not representative of all veterans – in particular, that the health of those in VHC is steadily worsening on average – a simple 'before' and 'after' analysis of spending for those now in VHC would produce biased results. That suggests that simplistic analysis is appealing but likely to be flawed.

Luckily, there is an incredibly rich set of data of the pre- and post-VHC experiences of individuals. This analysis uses information across a large number of individuals to compensate for the relatively short time period available, and allows the process of selection (including any systemic biases) into VHC to be considered. Or, in other words, the risks inherent in a lack of time in the data are addressed by using the depth of known individual experience including their access to the DP. We concentrate solely on gold card holders.

Various sophisticated econometric techniques were used in the analysis. The underlying model has two parts:

- An equation relating entry into VHC to the observed demographic characteristics of the individuals (age and gender) as well as their health care spending histories. Entry also depends on unobserved factors such as 'frailty' and 'illness'. As a proxy for this, the model has the capacity to include whether the person is in receipt of the DP and at what level.

- An equation explaining health care spending following entry into VHC.

An assessment of the available methods for programme evaluation led to a procedure in which the impact of VHC is estimated by comparing spending outcomes of individuals in VHC with spending outcomes of a group of matched non-VHC individuals. The matched non-VHC individuals were chosen because of their similar observed characteristics (gender, age, spending, DP status) to VHC individuals (immediately prior to the latter entering VHC).

This method does not simply compare spending outcomes for VHC and non-VHC individuals. That would introduce a bias if the unobserved factors are correlated with spending (for example, if the VHC individuals are 'sicker'). Rather it is based on changes in spending among both groups. For example, the VHC individuals may have higher spending both before and after they join VHC. The impact of VHC is measured through the extent to which spending changes. A rise in spending, relative to that of the matched non-VHC individuals, indicates a positive impact of VHC on spending.

1.6 Forecasts?

This is an evaluation model rather than a forecast model. It is therefore not suited to the latter purpose, in part because the modelling of changes undertaken here considerably simplifies the evaluation process, but complicates forecasting using this type of framework. That means that the analysis in this report says nothing about underlying trends in spending.

A simple forecast of VHC spending could be obtained by extrapolating past trends. But:

1) There are relatively few pre-VHC observations from which to model trends – about three years of data.
2) So any small change in trend could lead to a large change in the forecast.

Rather, forecasting is a more difficult exercise. It needs to account for:
1) Continuing volatility in private hospital spending.
2) The number of individuals entering the population of card holders.
3) The number of deaths.
4) Any rule changes.
5) Any changes in the behaviour of health service providers in the face of different rates inflation of costs and DVA reimbursements, etc.
6) A split into classes of individuals (males, females, ex-HACC, etc), with the forecasts of the individual classes aggregated into the total.

Obtaining accurate forecasts of overall spending in VHC, especially forecasts useful for estimating the historical impact of VHC, is therefore a large exercise.

1.7 Final thoughts

Although our comments have not been called for beyond an evaluation process, we take this chance to:
1) Commend the purchaser/provider split as good practice.
2) Recommend getting a better handle on veteran numbers still to transfer to VHC and the likelihood and rate of them doing so. Although the saving on an ex-HACC is larger than for a non-transitional, in part that is as they are already in a high care programme, and so they are costly members of the VHC population.
3) Recommend further analysis of the characteristics of veteran populations in different regions. That may also assist in determining whether underlying compositional differences go some way to explaining variations in unit costs. Such analysis of the acceptance rate would provide a useful foundation for effective benchmarking within VHC. It may also provide scope to benchmark against similar programmes such as HACC.

Any questions on this report should be directed to Andy Weiss or Anthony Baker on 6273-1222.

Access Economics
21 January 2003

2. INTRODUCTION

2.1 AIMS

This report seeks to determine whether VHC has resulted in cost reductions or increases since its introduction in January 2001.

2.2 EARLIER REPORT

The analysis flows on earlier work on the impact of VHC on spending across the first 10 months of the VHC programme (January 2001 to October 2001). The earlier report concluded there was higher spending following entry into VHC, but this finding was strongly qualified because:

- The programme was still in its early days.
- The time available for that report limited the range of econometric analysis.
- There were outliers in the available data.
- There was potentially important information not in the available data.

2.3 ADDITIONAL DATA

Data for an additional seven months is now available (to May 2002). DVA has commissioned Access Economics to assess the expanded dataset.

As at May 2002, there were 57,000 clients in VHC. Numbers were continuing to rise following an initial rush of clients into the programme, with the proportion of clients who have transferred over from the HACC scheme ('ex-HACCs') sitting at 27%.

2.4 APPROACH

Given that those in VHC are not representative of all veterans – in particular, that the health of those in VHC is steadily worsening on average – a simple 'before' and 'after' analysis of spending for those now in VHC would produce biased results. Accordingly, this study undertakes a detailed comparison of those in VHC with those outside VHC but closely 'matched' to those inside.

The resultant estimate of the change in spending is estimated directly first, and then allows for the possibility that those inside VHC are 'more frail' than the initially matched group outside.

2.5 ALLOWING FOR ALLIED HEALTH SAVINGS IN THE TRANSFER FROM HACC

VHC was established to "reduce the need for hospitalisation and other health care services, leading to savings in the veterans' health budget that will offset the cost of the program". VHC envisaged savings from two sources - veterans who were receiving HACC-like services from DVA and moved across to VHC; and veterans expected to come across from HACC.

DVA already delivered a range of HACC-like services to entitled veterans prior to the establishment of VHC. These services included respite care, assistance with aids and appliances, home modifications and certain community nursing and allied health services. DVA estimated that Veterans who received such services cost the HACC program and DVA combined some $3,500 annually per veteran. With the introduction of VHC, DVA estimated that the cost of providing these services would be $2,800 per veteran.

These savings are not the subject of this analysis, as they are not identifiable in the data. However, it is expected that a proportion of these savings would relate to allied health costs (which include community nursing) which are covered in this analysis.

2.6 Outline

The format of this report broadly follows the following layout:

- Section 3 discusses the approach taken to answer the central question: is VHC costing more or less?
- Section 4 details some initial results.
- Section 5 considerably extends the basic estimation by allowing for frailty effects.
- Section 6 gives detailed results by spending category and class of person.
- Section 7 asks if there is an underlying trend in the direction of savings over time.
- Section 8 interprets these results.
- Section 9 reports on the estimated effect of VHC on the number of service visits.
- Section 10 considers whether the methodology in this report could be used for forecasting.
- Section 11 offers some final thoughts.
- The technical appendix covers methodological issues in greater detail.

3. APPROACH TAKEN

This section covers the approach taken to address the central question: Has VHC added to or subtracted from DVA health spending over what otherwise would have been the case if it had not been introduced? The analysis uses the extra seven months of data now available.

The section begins with background in section 3.1. Next, section 3.2 briefly discusses the importance of getting the approach right. Section 3.3 discusses the choice of the model and associated methodology for estimating the impact of VHC. Section 3.4 gives a summary of the estimation method. It follows from the choice of methodology that a particular 'impact' is being estimated. Section 3.5 defines the impact. Section 3.6 highlights the data and modelling choices made in the analysis.

Details on the models and estimation methods, and many of the other issues discussed in the report, are given in the accompanying technical appendix.

3.1 BACKGROUND

One approach to assessing VHC impacts would be to use data from prior to the introduction of VHC to forecast aggregate spending since. Comparing the forecasts with actual spending would provide a simple assessment of the impact of VHC. However, we have not taken that approach. This is because:

1. There are relatively few pre-VHC observations from which to model trends (3 years of data), so a small change in trend unrelated to VHC could lead to a large change in the estimate of the VHC impact.

2. VHC has operated for a relatively short time and even with an additional 7 months of data, may not have yet fully settled into a steady-state.

3. Within aggregate spending there are two distinct patterns. First, VHC individuals have higher average spending than non-VHC individuals. Second, the spending of VHC individuals is increasing, whereas that for non-VHC individuals is relatively flat. Both suggest that VHC is 'cherry-picking' those who are more ill/costly clients. Failure to allow for that would overstate the cost of VHC.

Hence, the basic framework of the methodology for this stage of the analysis is as used for the previous stage. The analysis continues to:

1. *Make use of the individual record data*. As noted above, simplistic analysis is appealing but likely to be flawed. Luckily, there is an incredibly rich set of data of the pre- and post-VHC experiences of individuals. This analysis uses information across a large number of individuals to compensate for the relatively short time period

available, and allows the process of selection into VHC (including any systemic biases) to be considered. Or, in other words, the risks inherent in a lack of time in the data are addressed by using the depth of known individual experience.

2. *Employ comparative group analysis*. To make full use of the rich data available on individuals, the outcomes for VHC individuals are compared to non-VHC individuals. These comparisons are over individuals with similar characteristics.
3. *Undertake relative time analysis*. It is sensible to compare outcomes around the months in which individuals enter VHC. That means the analysis is in terms of 'months since entered VHC', rather than 'calendar months', as the latter would mix together people who had been in VHC for differing lengths of time.
4. *Allow for different usage across age and gender*. Age and gender are among the explanatory variables in the analysis. The different health care spending experience of men and women means that failure to allow for this could easily bias estimates of the net change in spending following the introduction of the VHC.

Given the additional time available for the consultancy, as well as the additional data, more analysis has been done. The analysis incorporates the new data and uses this additional information to estimate the overall saving or additional spending resulting from the introduction of VHC. A range of new methods have been employed and efforts have been made to test the sensitivity of the results to the methods used, the explanatory variables, and the choice of comparison group.

Note that at the time of the earlier report the data did not distinguish gold and white card holders. The new data does. The current analysis covers gold card holders only. That is important because they have around four times the spending of white card holders. Accordingly, one would expect comparably smaller changes to net spending for white card holders.

3.2 Getting the approach right

Getting the approach right involves issues such as:

- data quality,
- selecting the right comparator group, and
- assessing what is not provided by the available data – unobserved characteristics such as the overall state of health of each individual.

The methodology we use is the sum of the decisions we make to come to grips with these (and other) issues.

Data quality: Even with the range of methodologies investigated in this round, the analysis could fail if the underlying data are not up to scratch. One aspect of this is that the data should include relevant variables and should measure what it tries to measure – the data should be basically correct.

Comparator group selection: The estimation method should be appropriate for the data. Since the estimates are based on a comparative group analysis, some care must be taken in the choice and use of the comparative group – even if getting the perfect comparator group is impossible.

What we don't know: We assume that entry into VHC and health care spending can be modelled as a function of observed and unobserved characteristics of individuals.

Observed characteristics are those available in the data, such as gender, age, previous health care spending, card type, and a disability pension rate index.

Unobserved characteristics are, by definition, not available in the data. For example, the overall state of health of each individual is not available. The reasons for spending – illness, accident, etc – are not available. Amorphous characteristics such as 'frailty' are not available (though proxies for it might). Some of the unobserved characteristics may be relatively constant through time, such as overall health or 'frailty'; while others are more transitory, such as illness or minor accident.

The unobserved characteristics can influence entry into VHC as well as subsequent health care spending. That means that analysing health care spending independently of entry into VHC can lead to biased estimates of the impact of VHC.

To put it another way, say those who were 'more frail' were those selected to go into VHC. Naturally, spending for those in VHC would be higher and rise faster than for those 'less frail' outside the VHC. Failure to recognise that would lead to a biased estimator of the effect of VHC on health care spending.

3.3 Choice of Method

Within this general framework of observed and unobserved factors, a variety of models can be specified and estimation techniques applied. The methods fall into two general classes:

- **Those where detailed assumptions are made about the model up front:** The resulting estimators for these 'parametric methods' are preferable if the up front assumptions are correct. The Heckman selection estimator applied in our August 2002 report is in this class.
- **Those where fewer assumptions are made about the model up front:** 'Non-parametric methods' make fewer assumptions, and the estimators considered in this report are based on averages of the appropriate variables. Because the

methods do not make assumptions about the form of the model, they are potentially more robust.

In both approaches there is a model for entry into VHC. Entry into VHC is based on an underlying 'propensity', depending on the observed and unobserved factors. Put simply: the higher the propensity, the more likely an individual is to enter VHC. The methods differ in how they use the propensity in explaining health care spending.

In the Heckman estimator, the propensity feeds into a variable that allows for the correlation between entry and spending. In non-parametric methods, the propensity is used to carefully match the VHC individuals with non-VHC individuals with similar observed characteristics.

That means non-parametric methods take into account correlations between selection and subsequent health care spending resulting from *observed* variables. Conversely, and unlike the Heckman estimator, they do not explicitly model the correlation coming from the *unobserved* characteristics. Rather, they attempt to minimise its effect.

However, the non-parametric methods are also much more computationally intensive.

Our assessment, based on the experience gained from the analysis done for the initial report and the additional analysis done for this report, was that **a non-parametric method was the preferred approach**. The approach is summarised in the section 3.4. The time available for this consultancy also means that we were able to implement this method. [1]

3.4 Summary of Estimation Method

Our preferred method of estimating has been to use a matched difference-in-differences estimator. This sophisticated estimator includes both 'matching' and 'difference-in-differences'.

- **Matching** matches each VHC individual with one or more similar non-VHC individuals, and compares their spending. Thus, one comparison value is obtained for each VHC individual. Here, 'similar' means that the individuals are matched on their observed characteristics, as represented by their likelihoods for entering into VHC.
- **Difference-in-Differences (DID)** calculates changes ('differences') in spending for VHC individuals and non-VHC individuals between two time periods, one from before the VHC individuals enter VHC and one from after they enter. The method then compares the changes for VHC individuals with those for non-VHC individuals. Hence, it looks at the difference in the 'differences'.

[1] For example, obtaining basic estimates across 20 separate models – from males and females, ex-HACCs and non-transitionals, across five time periods – takes approximately 50 hours of computer time for each run.

- **Matched Difference-in-Differences** adds matching to the DID estimator. It therefore controls for observed characteristics determining entry into VHC. Again, one comparison value is obtained for each VHC individual.

The actual estimator averages the comparison values across the VHC individuals.

Matching compares the spending of VHC individuals with the spending of non-VHC individuals with similar characteristics. But estimators based only on matching are susceptible to sample selection bias if there are *unobserved* individual characteristics common to both entry into VHC and subsequent health care spending. For example, suppose that the VHC individuals are 'more frail' than the non-VHC individuals and subsequently have higher spending. The matching is likely to associate the higher VHC spending due to 'frailty' with the impact of VHC.

DID controls for unobserved individual characteristics that are constant over time. In the example just used, the VHC individuals have higher spending, both before and after they enter VHC. But because of the differences, all that matters is how the spending changes. Suppose VHC leads to a change in spending, whereas no change is expected for non-VHC individuals. The DID is the change for VHC minus the change for non-VHC. The latter is zero in this example.

The matched DID estimator combines matching and DID. It does not control for temporary unobserved factors. In other words, if there are temporary individual specific variables that affect entry into VHC and subsequent health care spending, then the estimator will be biased. The bias will depend on the nature of the temporary factor. For example, suppose that there is a death in the family and, as a result, the individual is given assistance through VHC. But subsequent health care spending rises. The increase is attributed to VHC. The same applies for a sudden rise in 'frailty' not yet reflected in pre-VHC entry health care spending, if the rise in 'frailty' leads to entry into VHC. (Think of a broken hip.)

The way in which the variables enter into the DIDs implies that a positive value of a DID corresponds to an increase in spending as a result of VHC.

3.5 What 'impact' is being estimated

The non-parametric methods are based on differences in average spending. Hence, the impact of VHC is estimated in the same way. The matching of VHC individuals with similar non-VHC individuals implies that the procedure is asking: What would have happened to the spending of individuals chosen for VHC had they not actually received any home care? The impact is known as *treatment on the treated*.

The way in which the impact is estimated means that features such as policy changes, seasonal effects, and inflation that affect all individuals in the same way will not influence the estimates and can effectively be ignored.

Note that *treatment on the treated* differs subtly from the impact estimated in our earlier report, which asked what might have happened to the spending of VHC individuals had they not been chosen for VHC.

3.6 DATA AND MODELLING CHOICES

3.6.1 Spending categories

Seven categories are modelled to determine the impact of VHC on health spending:

- **Private hospitals (PH)** covers mainly the accommodation costs of veterans, with billing direct to DVA.
- **General practitioners (GP)** covers GP consultations.
- **Specialists (SC)** covers specialist consultations either in hospital or private rooms.
- **Diagnostic imaging (DI)** covers X-ray and ultra-sound type activities such as magnetic resonance imaging.
- **Procedures (PR)** covers a range of operations everything from open heart surgery to the removal of an ingrown toenail performed by a GP or Specialist either in hospital or private room.
- **Allied health (AH)** covers activities from community nursing, occupational health, speech pathology, social work and podiatry.
- **Pathology (PA)** covers pathology and a range of miscellaneous activities.

The analysis estimates the impact of VHC on each category. The total impact is obtained by basically aggregating across the seven categories. We also investigate the impact of VHC on the *number* of allied health, general practitioner and specialist services. The way in which services in the other categories are billed, as evidenced by the raw data, suggests that it would be meaningless to model the number of services in those categories.

3.6.2 Basic sample

The basic sample includes all VHC individuals and a ¼-sized sample of the non-VHC individuals (more formally, each non-VHC individual is allowed to enter the sample with probability ¼).

Individuals less than 50 years old in January 2001 are excluded, as are individuals with missing data.

3.6.3 Classes of individuals

There are significant differences in the characteristics of males and females entering VHC, and also between those who are 'ex-HACC' and those who are 'non-transitionals'.

Hence, we separately estimate the impact on the four classes – male ex-HACC, female ex-HACC, male non-transitional, and female non-transitional.

3.6.4 Decision and treatment months

We classify VHC individuals by the month in which they enter VHC. The spending comparisons are based on this classification. For example, in the matched DID estimator, we pick out individuals who entered VHC in a particular month. We estimate how their spending changed over a subsequent period. We then do the same for the matched non-VHC individuals, using the same set of months.

The 'particular month' in the previous paragraph is called the *decision month*. In this analysis, we focus on the following decision months: March 2001, June 2001, September 2001, December 2001 and March 2002. Thus, for individuals entering VHC in March 2001, for example, there are fourteen months of post-entry data. For individuals entering in March 2002, there are only two months of post-entry data. The sample of months is used to speed the computations, while still covering the period in which VHC has been operating.

The months after the decision month are referred to as the *treatment months*. The treatment months run from the month after the decision month up to and including May 2002.

3.6.5 Comparison groups and explanatory variables

In practice, the matching in the matched DID estimation is done on the basis of estimated propensity scores, where the propensity score is the likelihood of entry into VHC given the observed characteristics, rather than on a set of explanatory variables. Matching on large set of variables is very data and computation demanding. The scores are obtained from estimated models known as probit models (where the dependent variable is reduced to a zero-one format, with the 'one' for the individuals entering into VHC in the decision month and the 'zeroes' for the non-VHC individuals in the comparison group).

The explanatory variables include age, state of residence, and health care spending in the two months prior to the decision month. Different models are estimated for the four classes of individuals – male ex-HACC, female ex-HACC, male non-transitional, and female non-transitional.

The comparison group is also used as the pool of non-VHC individuals from whom the matched individuals are found.

For each decision month, we form the comparison group by taking a random sample from the set of all individuals who were not in VHC as at the decision month and who did not enter VHC in the five months following the decision month. For each decision month the sample contains approximately 7,000 individuals (males and females).

3.6.6 Base period for spending comparisons

Spending tends to increase in the months prior to entry into VHC. Alternatively, the higher spending could be seen as associated with, or leading to, the entry into VHC. The matching implies the matched non-VHC individuals have similar spending patterns, as represented by the propensity scores, although the patterns will not be exactly the same. The base period for the DID comparisons is chosen to reduced the effects on the DID estimates of any differences in the spending patterns.

As noted in the previous section, the two months prior to the decision month are included in the propensity score models. Given these two months, prior months are generally not significant.

In forming the base period for the DID comparisons, we:

- *Exclude the decision month*. It is not clear whether to attribute the spending in this month to the pre-VHC period or the post-VHC;

- *Exclude the two months prior to the decision month*. Spending is often higher in these months and it is not appropriate to assign any subsequent fall in spending to VHC; but,

- *Include months 3 to 8 prior to the decision month*. The average spending over those 6 months is therefore the base amount used in the comparisons.

For example, for decision month December 2001:

- December 2001 is excluded from the analysis.
- October and November 2001 are included in the propensity model but excluded from the base period.
- May 2001 to September 2001 are used as the base for the DID comparisons. The DID estimates are based on changes between the base and January 2002, the base and February 2002, etc.

If spending is higher just prior to the decision month, then it may also be higher immediately after the decision month. The interpretation of the results should take this into account.

The choice of the length of the base period is associated with the problem of outliers, and is discussed further in the next section.

3.6.7 Outliers and trimmed means

As noted in the August 2002 report, there are small numbers of individuals with large spending in some categories. That is, there are data outliers. The matched DID estimators are based on averaging and so reduce the impact of outliers. But in some categories of spending, especially private hospitals, the outliers are extreme and just a handful of observations can potentially dominate the results.

For a particular treatment month, the contribution towards the matched DID estimate from a particular VHC individual is the *difference* between the individual's spending in that month and the individual's average spending over the six months in the base period (see the example in section 3.6.6). Hence,

difference = spending in treatment month – average monthly spending in base period

The contribution from the matched individuals is the average *difference* over the matched individuals. The DID is

DID = *difference* for VHC individual – average *difference* for matched individuals

The matched DID estimator is obtained by averaging the DIDs in the last equation.

Of the individuals contributing to the DID, an extreme DID is more likely to arise because of an extreme difference for the VHC individual, since the second term on the right hand side of the last equation is typically an average over a reasonable number of individuals (so large positive and large negative values cancel each other out). High spending for the VHC individual in the base period relative to the treatment month implies a large negative value of the *difference* – a contribution to an estimate implying a cost saving. High spending in the treatment period implies a large positive *difference*.

One way to check the importance of outliers is to compare the matched DID estimator with a trimmed matched DID estimator. The latter leaves out observations for which the difference (either as a percentage or in absolute dollars) is large.

An alternative is to increase the length of the base period to directly reduce outlier risks. There is no optimal value for the number of months in the base period. The end of the base period (the month closest to the decision month) is chosen as discussed above – we exclude the months included in the propensity model and which are most influenced by any pre-entry increase in spending. At the other end, going back too far implies that the pre-entry health conditions of the individuals are not properly represented. Not going back far enough exposes the estimation to increased risks of volatility due to outliers.

We have adopted a six month base period. That gives rather less volatile results than does a three month base period.

3.6.8 Private hospital spending

As noted in the previous section, the outliers are most severe in the private hospital category.

The importance of private hospital spending in total spending suggests that additional checks should be conducted. However, the issue is not simple. For example:

- For any decision month (and relative to individuals not entering VHC), individuals entering VHC have, on average, higher spending just prior to entry.
- Only a small percentage of individuals have private hospital spending in a particular month, and private hospital spending does not guarantee entry into VHC.
- Individuals with private hospital spending in a particular month have a higher probability of having private hospital spending in future months.
- Private hospital spending is just one of many variables in the propensity models. While the propensity model implies that individuals are matched with others with similar propensity scores, it is not necessary that they be matched with someone with similar private hospital spending.

In the August 2002 report, we deleted observations with private hospital spending over $4,000 a month. In the current analysis, private hospitals outliers are addressed by averaging spending across a six month base period, which allows for distortions from big spending to be smoothed by a number of other months.

3.6.9 Additional data

The postcode of each individual, at the time of spending, is available in the dataset. These were mapped into DVA regions via postcode-SLA and SLA-region concordances. Some observations were lost due to recent postcode changes and mismatches in the two concordances – the SLA-region concordance pre-dates some recent changes to SLA definitions. Postcodes were also mapped into State/Territory codes and codes for capital cities, other metropolitan centres, large rural centres, small rural centres, other rural areas, remote centres, and other remote areas.

The disability pension rate for each individual for each month is also available. These were mapped into five categories: persons without service related disabilities, low disability (5-95% of the general rate), medium (100% of the general rate), special disability (special rate and Intermediate rate combined) and extreme (extreme disability adjustment).

The special disability category is detailed below.

I. Special rate
- Totally permanent incapacity
- Blinded
- Temporarily totally incapacitated
- CLC class C (old code for blinded)

II. Intermediate rate
- Intermediate
- CLB class B (old code for intermediate)

3.6.10 Other issues

Other issues include:

1. the choice of variables for the models,
2. the sensitivity of the results to particular assumptions, and
3. the changes in the models and the characteristics of the individuals through time.

4. INITIAL ESTIMATES

4.1 INTRODUCTION

In this section of the report, we give the initial estimates of the impact of VHC on health care spending – that is, those before we attempt to adjust for the impact of 'frailty'. We give the results when the new disability pension rate variable is not included in the analysis (section 4.2). We then add the latter variable as a proxy for 'frailty' to test its significance to VHC spending (section 5).

This serves to highlight the role of 'frailty' in the analysis.

A proper interpretation of the results depends on more than just the overall estimates. It also depends on the estimates across classes of individuals, categories of spending, and decision and treatment months. More detailed results are given in section 6, while a discussion of whether the results point to an improving trend over time is in section 7. A full interpretation of the results is given in section 8.

A proper interpretation of the results also depends on the caveats to the results. Caveats are given in section 5.1 and are discussed further in section 8.

4.2 OVERALL RESULTS

In brief, in the basic model and combining over individuals, categories, and months, it is estimated that VHC individuals have, on average, $15.10 per month lower spending in the treatment months than do comparable non-VHC individuals. In other words VHC is associated with a decrease in spending of $15.10 per person per month.

This compares with average monthly spending of individuals in VHC of approximately $600 (that is, it is the equivalent of 2.5%) and average spending of non-VHC individuals of just under $300.

Within this overall result, there are complicated patterns of spending across classes of individuals, categories of spending, decision and treatment months. These patterns, as well as the overall results, are detailed in section 6 of the report. Here, we concentrate on relatively aggregated results. We note that:

1) The fall in spending is not uniform across the seven categories. Generally, there is a spending increase for allied health and GPs and a spending decrease for private hospital and procedures.
2) But private hospitals and procedures are also the most volatile elements of spending.

The relative pattern of spending noted in point 1) has important implications for the interpretation of the results – whether the estimated impact of VHC can be viewed:

- as reflecting the preventative aims of the programme, as might be the case with an increase in allied health and GP spending and a fall in private hospital spending; or
- as dominated by a change in private hospital spending that is too volatile to allow any statistical confidence to be attached to its estimates.

That is, point 2) is one of the main caveats to the results. Other caveats are given in section 5.1.

5. 'FRAILTY'

It is evident from the discussion that the unobserved concept loosely defined as 'frailty' might have a potentially important effect on the estimates. The matching in the matched DID estimator is necessarily based on observed variables, and if 'frailty' is an important determinant of post-VHC spending then not matching on it can lead to either biased estimators or larger standard errors.

Access Economics has been provided with the disability pension rates assessed for the individuals. The actual rates were transformed into a disability variable with five classes (zero, low, moderate, high, and extreme), and the disability variable was added into the analysis.

The hypothesis tested here is whether the disability variable contains information on 'frailty' beyond that in the existing variables – age and previous health care spending.

The disability pension rates are for disabilities resulting from war service, and so there are relatively few females with non-zero rates. In contrast, approximately two thirds of males have non-zero values. Hence, the main effects of the variable are likely to be seen in the estimates for males.

Explanatory variables enter into the analysis via the propensity score models. Re-estimation of the latter showed that the disability variable is indeed a significant determinant of entry into VHC – allowing for the extra information in the disability variable improves the ability to predict whether someone went into VHC or not.

The variable may be reflecting true 'frailty', or it may represent a piece of information used in the assessment process (irrespective of 'frailty'). In any case, the disability variable has an impact on the matching.

In the expanded model, combining over individuals, categories, and months gives a decrease in spending of $21.32 per person per month. Thus, the estimate of savings is larger when disability is taken into account.

The larger value can be explained as follows. In the matched DID estimator, the contributions from the VHC individuals are unchanged when the disability variable is added to the propensity score models. But the contributions from the non-VHC individuals change.

For example, the matching may pick out different non-VHC individuals, who have more comparable levels of disability. Consider a highly disabled VHC veteran. Suppose that the non-VHC veterans now matched to highly disabled VHC veteran have higher levels of disability than those in the model without the disability variable.

Because of the higher levels of disability of the 'new' matched non-VHC individuals, their spending may increase more through time than did the spending of the 'old' matched non-VHC veterans. The larger increase for the non-VHC individuals leads to a lower value of the DID and a higher estimate of the saving associated with VHC.

Chart 4 shows that allowing for 'frailty' has an impact exactly where it would be expected to – on men rather than women (as by and large it is men who receive Disability Pensions), and on non-transitionals rather than ex-HACCs (HACC is a similar program to VHC, so those who are 'more frail' may well have already gone into HACC prior to their transfer into VHC).

Note that the Disability Pension therefore seems to work well as a proxy for 'frailty' for men. The lack of an equivalent proxy for women suggests that these results may understate the true savings under VHC.

Chart 4 – Effect of allowing for 'frailty' by class of person

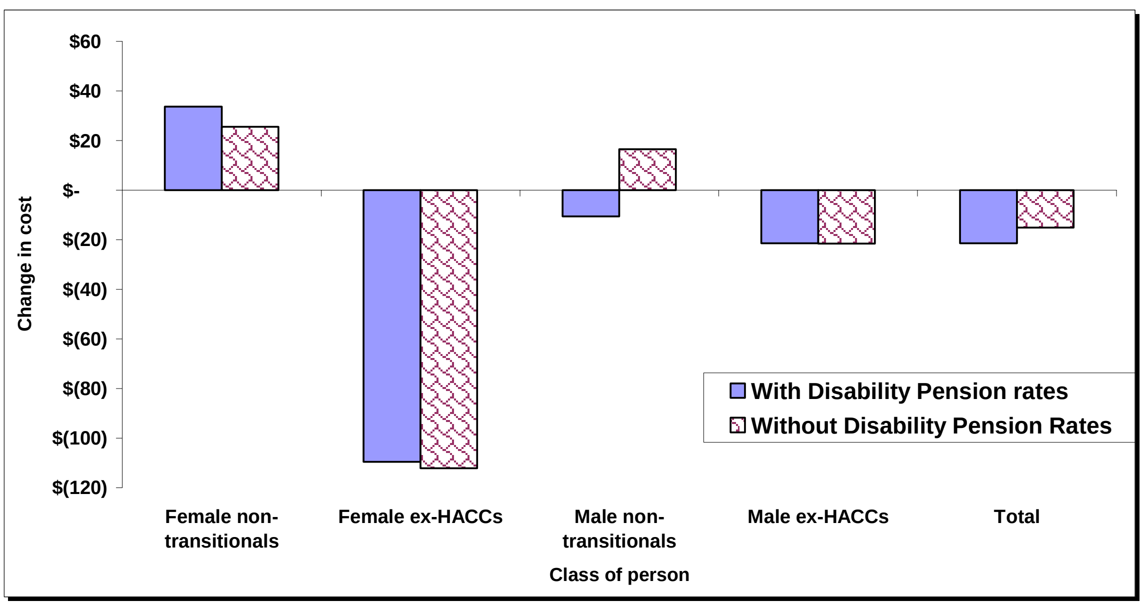
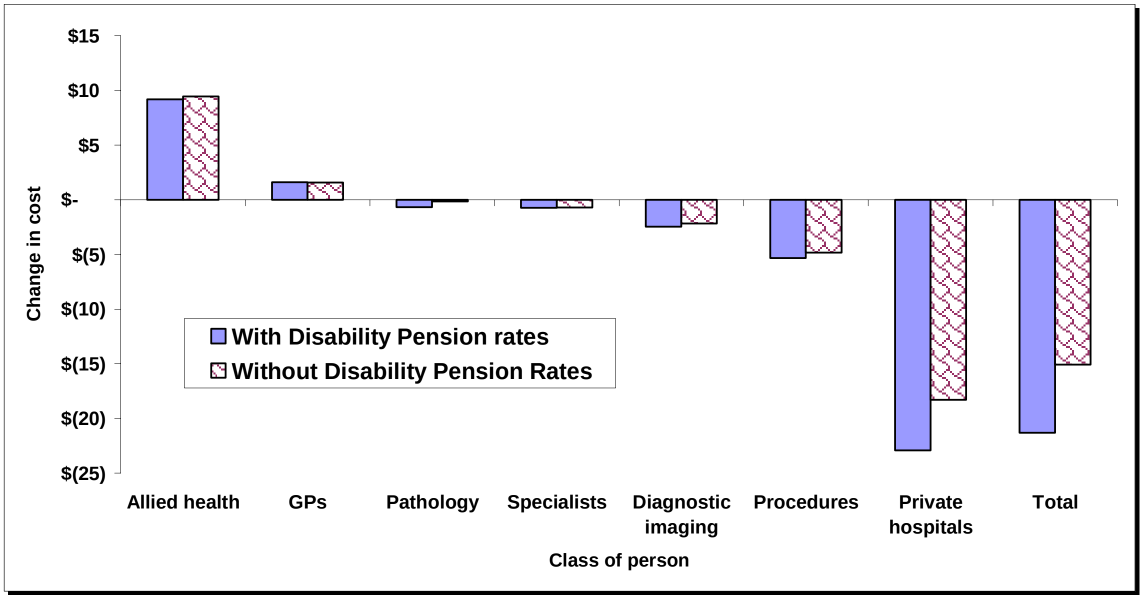


Chart 5 does not necessarily strengthen or weaken the case that using the Disability Pension as a proxy for 'frailty' was correct. It does, however, make the basic point that it is spending on private hospitals which is most volatile.

Chart 5 – Effect of allowing for 'frailty' by category of spending



Of course, the caveats on the results still hold.

5.1 CAVEATS

Just as in our August 2002 report, a set of general caveats should be noted. Indeed, although this report has involved rather more sophisticated modelling techniques and a longer run of data, the caveats are similar to those in our earlier report. Some are generic to statistical modelling.

Because a different estimation methodology is applied, the particulars of the caveats change:

- **Potentially important information may not be in the available data**. While the data covers a longer period, it contains few extra variables. Unobserved effects such as 'frailty' and 'illness' remain unobserved. The report has described how the methodology allows for some unobserved factors, but may be biased if others are present. This section addresses this using a proxy for 'frailty' in the analysis, but there is no guarantee that the results are still unaffected by other unobserved variables.
- **The results are sensitive to the treatment of outliers**. Outliers are handled by using a six month base period (thereby smoothing some volatility in the estimates). However, that does not remove the underlying difficulty that outliers still generate much of the volatility in spending, and so any method of dealing with them runs the risk of throwing away useful information.
- **It is 'one-size-fits-all' modelling**. A total of 20 different propensity models are estimated, covering gender, transitional status (ex-HACC or not), and time. The

same set of variables is used in each. No attempt was made to delete insignificant variables from the models. No other model is used in the matched DID estimator, although the assumptions required for the estimator to be unbiased were noted.

- **Error estimates are probably too small**. Obtaining theoretically correct estimates of standard errors in matched DID estimators is difficult. As before, the estimated standard errors do not take into account all potential sources of uncertainty, and so are probably too small. The implication is that the results are probably less definitive than they may otherwise seem.

6. DETAILED RESULTS

For equivalent individuals, spending is estimated to be lower in VHC than for a matched group of veterans outside VHC, by an average of $21.32 per VHC participant per month.

This compares with the average monthly health cost of individuals in VHC (of approximately $600 per month) and of non-VHC individuals (under $300 per month). It therefore implies a saving of 3.6% for those inside VHC.

1. The evidence for a net saving is strongest for female ex-HACCs, and weakest for female non-transitionals.

2. Most of the savings are driven by falls in private hospitals and procedures spending. That is no surprise, given that these are the most expensive categories, and the most volatile. Against the trend, spending in preventative areas – allied health and general practitioners – tend to increase, at least in the short term.

3. Once allowance for the disability proxy is made, there is some evidence that earlier entrants have demonstrated a greater monthly cost saving than later entrants. Evidence suggests that this pattern, rather than being due to initial entrants being more frail, may be due to the VHC program causing a short term (preventative?) increase, but a sustained long term decrease in health costs.

Table 3 – Cross-tabulation of average spending changes

	Female	Male	Total
Non-transitional	+ $ 33.70	- $ 10.52	+ $ 9.50
Ex-HACC	- $ 109.68	- $ 21.49	- $ 70.27
Total	- $ 28.64	- $ 14.25	- $ 21.32

6.1 RESULTS BY CATEGORY

Table 4 (and its corresponding Chart 6) displays the breakdown of the overall net cost savings (or increases) by category of health spending, grouped by class of person.

A number of features are apparent in the results:

1. There is a difference between classes of people – between males and females, and between ex-HACCs and non-transitionals. There are greater savings for women, and greater savings for ex-HACCs. The latter is somewhat counterintuitive – although VHC doubtless tries to do better than HACC, in many ways it contains elements of a HACC-like programme.

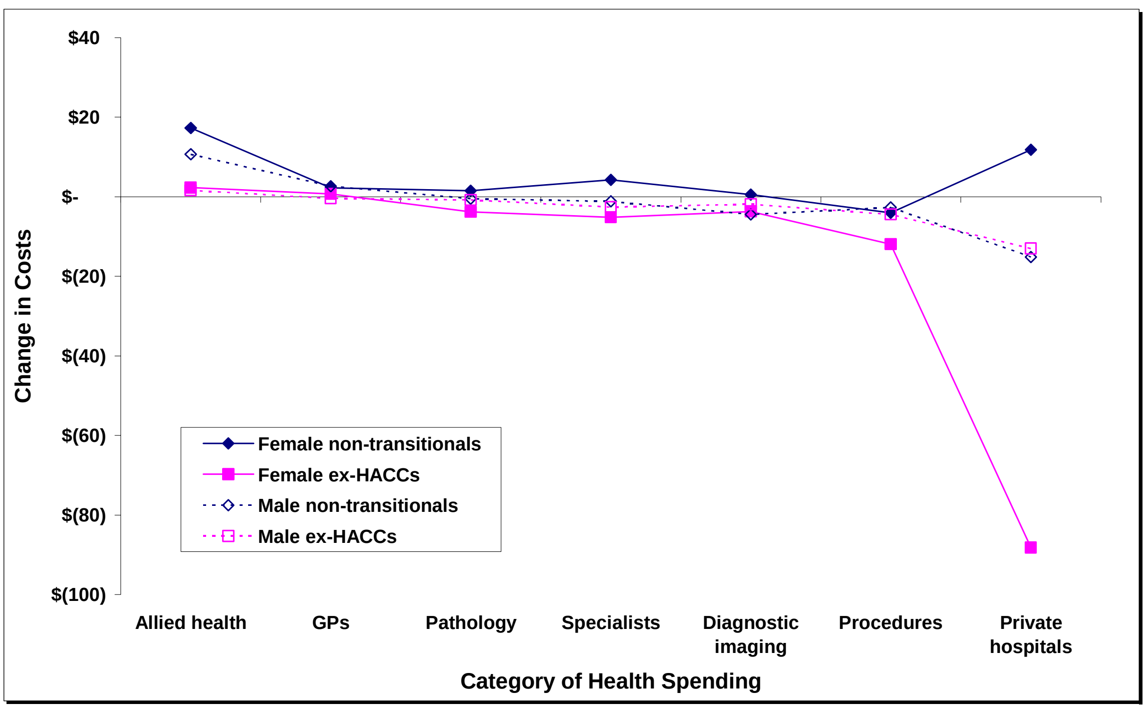
2. Private hospitals dominate the results, being the greatest contributor to the bottom line, as well as generating the largest volatility between classes.

3. Allied health and GP spending go against the trend, registering an increase, especially for non-transitionals.

Table 4 – Detail by gender and transitional status

Gender	Non-transitional or ex-HACC	Number	Allied health	GPs	Pathology	Specialists	Diagnostic imaging	Procedures	Private hospitals	Total
Fem.	Non-T	4304	+17.33	+2.23	+1.54	+4.28	+0.56	-4.03	+11.79	+33.70
Fem.	Ex-H	3311	+2.29	+0.74	-3.76	-5.13	-3.68	-11.93	-88.22	-109.68
Male	Non-T	5203	+10.69	+2.67	-0.45	-1.15	-4.44	-2.68	-15.16	-10.52
Male	Ex-H	2675	+1.63	-0.40	-0.81	-2.58	-1.88	-4.43	-13.00	-21.49
Overall		15493	+9.17	+1.61	-0.67	-0.74	-2.44	-5.33	-22.92	-21.32

Chart 6 – Change in Costs by Category



6.2 Results by month of entry

Table 5 (and Chart 7) displays the breakdown of the overall net cost savings (or increases) by category of health spending, grouped by month of entry.

As before, private hospitals dominate, and allied health shows a cost increase. The most prominent feature is that, following allowance for a proxy for 'frailty', there is a time trend – earlier entrants exhibit overall cost savings, and later entrants exhibit cost increases. Four possible explanations of this are:

5. Earlier cohorts (those people who joined the VHC programme earlier) were different in some significant way, such as being more or less frail, or having been 'cherry-picked'. Most of the HACC transferees were moved in June and July of 2001, and in the months leading up to it, and that may affect the results.

6. We have more data for earlier cohorts – and there is a short term cost increase but a long term cost saving (which is only apparent for earlier cohorts).

7. We have more data for earlier cohorts – but our numbers for later cohorts have less confidence associated with them (that is, larger standard errors). For example, individual monthly spending on private hospitals can range up to $110,000. At the time of our analysis, we had only two months of reliable spending data for 2002-March entrants, yet earlier cohorts had up to 14 months of data.

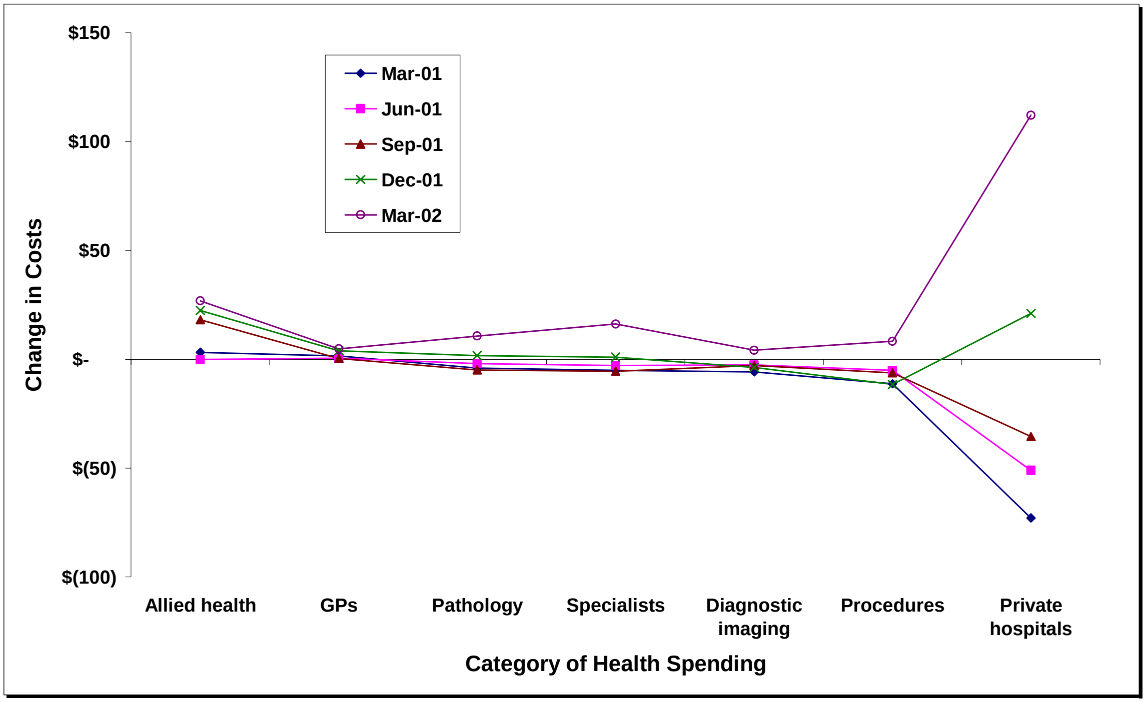
8. It is the result of chance, with the differences across cohorts essentially driven by the volatile private hospital category (see Chart 7).

This issue is examined in more detail in the next chapter, on time trends.

Table 5 – Detail by Month of Entry

Month they joined VHC	Number	Allied Health	GPs	Pathology	Specialists	Diagnostic Imaging	Procedures	Private Hospitals	Total
2001-Mar	2389	+3.17	+1.54	-3.95	-5.02	-5.79	-11.22	-72.88	**-94.17**
2001-Jun	7019	-0.10	+0.51	-1.98	-2.91	-2.63	-5.10	-51.01	**-63.20**
2001-Sep	2248	+18.11	+0.45	-4.83	-5.49	-2.89	-6.23	-35.41	**-36.27**
2001-Dec	1904	+22.42	+3.83	+1.67	+1.00	-3.73	-11.59	+21.10	**+34.69**
2002-Mar	1933	+26.83	+4.81	+10.68	+16.22	+4.16	+8.27	+112.02	**+182.99**
Overall	15493	+9.17	+1.61	-0.67	-0.74	-2.44	-5.33	-22.92	-21.32

Chart 7 – Change in Costs by Month of Entry



6.3 Correlations between categories

Intuitively, some categories of health spending are associated with others. For example, GPs will refer people to specialists, and those undergoing procedures may stay in hospital around that time.

Table 4 shows that allied health and general practitioner spending goes up, on average, and that the other five categories exhibit savings. The correlations (a measure of association) between spending by categories are given in Table 6. This table suggests a division of the seven categories of spending into two broad groups:

1. Allied health and general practitioners (broadly a 'prevention' category).
2. Pathology, specialists, diagnostic imaging, procedures and private hospitals (dominated by the dollars of a broadly 'treatment' or 'cure' category).

Table 6 – Cross-correlations of spending change, by category

	AH	GP	PA	SC	DI	PR	PH
Allied health	100%	13%	5%	5%	3%	3%	6%
GPs	13%	100%	19%	15%	16%	6%	17%
Pathology	5%	19%	100%	56%	51%	54%	52%
Specialists	5%	15%	56%	100%	44%	32%	56%
Diagnostic imaging	3%	16%	51%	44%	100%	37%	38%
Procedures	3%	6%	54%	32%	37%	100%	61%
Private hospitals	6%	17%	52%	56%	38%	61%	100%

There is relatively little association between spending in the categories in the first group (allied health and GPs) and spending in the categories in the second group (specialists, pathology, diagnostic imaging, procedures and private hospitals).

In other words, a trip to an allied health professional or a GP did not usually end up also being associated with a visit to a specialist or any tests (such as pathology or diagnostic imaging), procedures and private hospital stays.

(As an aside, it is not surprising that GP spending is more highly correlated with the categories in the second group than is allied health. For example, GPs visit sick patients in hospitals.)

Furthermore, spending in the categories in the second group is more closely associated with each other than they are to spending in the first group.

That broad grouping may well be consistent with a general split of these categories into 'prevention' (allied health and GPs) versus 'cure' or 'treatment' (the others). These results are potentially consistent with the hypothesis of an increase in short term preventative spending, which leads to a longer term reduction in curative spending.

6.4 Confidence

As well as an overall estimate (a $21.32 monthly saving), an inspection of the standard errors gives an insight into the confidence we can put into these findings, or how clearly the trend stands out from inherent differences in the sample.

The standard error of the mean is the standard deviation divided by the square root of the number of observations, and the t-score is the ratio of the point estimate to the standard error.

The generally accepted informal rule is that a t-score above 2 or below -2 indicates a significant result (a result in which one can be relatively more confident).

Table 7 details the standard errors and t-scores of the estimated average cost change, by class of person and category of health spending.

Table 7 – Standard Errors of Estimates

	Female non-transitionals			Female ex-HACCs		
	Estimate	**Std Err**	**t-score**	**Estimate**	**Std Err**	**t-score**
Total	$33.70	$22.74	1.48	-$109.68	$17.84	-6.15
Allied health	$17.33	$2.46	7.05	$2.29	$1.34	1.70
GPs	$2.23	$0.69	3.23	$0.74	$0.66	1.12
Pathology	$1.54	$1.13	1.35	-$3.76	$0.82	-4.57
Specialists	$4.28	$1.35	3.16	-$5.13	$1.16	-4.44
Diagnostic imaging	$0.56	$1.39	0.41	-$3.68	$1.29	-2.86
Procedures	-$4.03	$3.08	-1.31	-$11.93	$2.61	-4.57
Private hospital	$11.79	$17.70	0.67	-$88.22	$13.54	-6.51
# Observations	4304			3311		

	Male non-transitionals			Male ex-HACCs		
	Estimate	**Std Err**	**t-score**	**Estimate**	**Std Err**	**t-score**
Total	-$10.52	$21.95	-0.48	-$21.49	$23.24	-0.92
Allied health	$10.69	$2.45	4.36	$1.63	$1.99	0.82
GPs	$2.67	$0.70	3.83	-$0.40	$0.73	-0.55
Pathology	-$0.45	$1.10	-0.41	-$0.81	$1.21	-0.67
Specialists	-$1.15	$1.26	-0.92	-$2.58	$1.48	-1.74
Diagnostic imaging	-$4.44	$1.48	-3.00	-$1.88	$1.70	-1.11
Procedures	-$2.68	$3.15	-0.85	-$4.43	$3.76	-1.18
Private hospital	-$15.16	$17.02	-0.89	-$13.00	$17.23	-0.75
# Observations	5203			2675		

A first inspection concludes that, apart from female ex-HACCs, the total cost saving estimates for VHC are not significant. However, this number is dominated by private hospital spending, which is quite volatile. As noted above, monthly costs can range above $100,000, but are often zero. The statistical procedure which estimates 'what would have happened without VHC' therefore naturally yields a larger estimate of error.

Among spending increases, the increase in allied health and GP spending has high t-scores for non-transitionals, and female ex-HACCs demonstrate significant savings in all other categories.

That analysis suggests that the most stable element in the results is a small increase in allied health and GP spending and that, in general, there are notable limits on the confidence to be attached to the average estimated saving inside VHC of $21.32 per client per month.

6.5 Detailed listing

Table 8 depicts the estimated cost savings, broken down by gender, transitional status, month of entry, and category of health spending.

Table 8 – Detailed Cost Impact

Gender	Non-trans. or ex-HACC	Month they joined VHC	Number	Allied health	GPs	Pathology	Specialists	Diagnostic imaging	Procedures	Private hospitals	Total
Overall			**15493**	**+$9.17**	**+$1.61**	**-$0.67**	**-$0.74**	**-$2.44**	**-$5.33**	**-$22.92**	**-$21.32**
Summaries											
Fem.			7615	+$10.79	+$1.59	-$0.77	+$0.19	-$1.28	-$7.46	-$31.69	-$28.64
Male			7878	+$7.61	+$1.63	-$0.57	-$1.64	-$3.57	-$3.28	-$14.43	-$14.25
	Non-T		9507	+$13.70	+$2.47	+$0.45	+$1.31	-$2.17	-$3.29	-$2.96	+$9.50
	Ex-H		5986	+$1.99	+$0.23	-$2.44	-$3.99	-$2.88	-$8.58	-$54.60	-$70.27
		2001-Mar	2389	+$3.17	+$1.54	-$3.95	-$5.02	-$5.79	-$11.22	-$72.88	-$94.17
		2001-Jun	7019	-$0.10	+$0.51	-$1.98	-$2.91	-$2.63	-$5.10	-$51.01	-$63.20
		2001-Sep	2248	+$18.11	+$0.45	-$4.83	-$5.49	-$2.89	-$6.23	-$35.41	-$36.27
		2001-Dec	1904	+$22.42	+$3.83	+$1.67	+$1.00	-$3.73	-$11.59	+$21.10	+$34.69
		2002-Mar	1933	+$26.83	+$4.81	+$10.68	+$16.22	+$4.16	+$8.27	+$112.02	+$182.99
Details											
Fem.	Non-T	2001-Mar	570	+$10.62	+$4.07	+$1.60	-$3.08	-$3.95	-$8.13	-$71.50	-$70.37
Fem.	Non-T	2001-Jun	1276	+$4.83	+$2.44	-$4.68	-$0.48	-$1.90	-$8.02	-$64.34	-$72.16
Fem.	Non-T	2001-Sep	948	+$23.25	-$0.43	-$0.74	-$2.01	+$2.82	-$6.14	+$9.15	+$25.91
Fem.	Non-T	2001-Dec	733	+$20.03	+$2.53	+$0.89	+$1.17	-$10.07	-$19.86	-$1.34	-$6.65
Fem.	Non-T	2002-Mar	777	+$32.99	+$3.52	+$15.09	+$28.09	+$15.20	+$23.06	+$213.52	+$331.47
		Overall	**4304**	**+$17.33**	**+$2.23**	**+$1.54**	**+$4.28**	**+$0.56**	**-$4.03**	**+$11.79**	**+$33.70**
Fem.	Ex-H	2001-Mar	577	-$2.68	-$1.52	-$7.34	-$9.74	-$6.18	-$22.93	-$123.46	-$173.84
Fem.	Ex-H	2001-Jun	2351	+$2.62	+$1.12	-$2.96	-$3.01	-$1.76	-$6.32	-$77.34	-$87.66
Fem.	Ex-H	2001-Sep	115	-$4.86	+$3.16	-$2.41	-$15.63	-$12.84	-$19.67	-$67.42	-$119.67
Fem.	Ex-H	2001-Dec	145	+$12.63	+$3.38	-$8.74	-$13.83	-$8.17	-$50.04	-$124.76	-$189.52
Fem.	Ex-H	2002-Mar	123	+$13.78	-$1.20	+$2.16	-$3.89	-$14.66	-$15.39	-$107.13	-$126.32
		Overall	**3311**	**+$2.29**	**+$0.74**	**-$3.76**	**-$5.13**	**-$3.68**	**-$11.93**	**-$88.22**	**-$109.68**
Male	Non-T	2001-Mar	828	+$3.14	+$3.06	-$4.94	-$3.05	-$8.95	-$10.14	-$52.32	-$73.19
Male	Non-T	2001-Jun	1475	-$5.34	+$0.15	-$0.89	-$5.73	-$4.59	-$0.06	-$48.06	-$64.53
Male	Non-T	2001-Sep	1071	+$15.33	+$0.00	-$9.44	-$7.61	-$7.14	-$6.46	-$67.89	-$83.21
Male	Non-T	2001-Dec	886	+$23.70	+$5.52	+$4.28	+$3.15	-$0.43	+$0.22	+$46.50	+$82.93
Male	Non-T	2002-Mar	943	+$24.92	+$6.64	+$9.92	+$10.96	-$0.93	+$1.34	+$70.87	+$123.72
		Overall	**5203**	**+$10.69**	**+$2.67**	**-$0.45**	**-$1.15**	**-$4.44**	**-$2.68**	**-$15.16**	**-$10.52**
Male	Ex-H	2001-Mar	414	+$1.09	-$0.72	-$4.92	-$5.08	-$1.48	-$1.33	-$45.41	-$57.85
Male	Ex-H	2001-Jun	1917	-$2.67	-$1.23	+$0.18	-$2.22	-$2.68	-$5.51	-$12.11	-$26.23
Male	Ex-H	2001-Sep	114	+$24.69	+$9.30	+$2.11	-$4.28	-$0.45	+$8.84	-$68.40	-$28.19
Male	Ex-H	2001-Dec	140	+$36.95	+$0.51	-$0.03	+$1.87	+$13.08	-$3.21	+$128.87	+$178.05
Male	Ex-H	2002-Mar	90	+$11.47	+$4.88	-$7.81	-$3.66	-$11.99	-$14.42	-$33.55	-$55.09
		Overall	**2675**	**+$1.63**	**-$0.40**	**-$0.81**	**-$2.58**	**-$1.88**	**-$4.43**	**-$13.00**	**-$21.49**

7. TIME TRENDS

Allowing a proxy for 'frailty' in the estimation also raised up an interesting possibility, examined further here, where we look at month-by-month estimated cost changes, relative to the time the participant entered the VHC program.

For example, monthly estimates for April and May 2001, for someone who entered in March 2001, are compared to the monthly estimates for January and February 2002, for someone who entered in December 2001.

The data finishes in May 2002 for all individuals. As a result, we have less data for those who entered later, and consequently see a greater spread (increasing volatility and decreasing reliability) in our estimates the further out from the decision month we look.

Nonetheless, the following two charts show that:

1. Spending in allied health and GPs (the 'preventative' categories) are initially higher, but settle in the long term to a net effect of around zero. Later charts show that this effect is more noticeable for non-transitionals.

2. Spending in the other 'curative' categories is initially unaffected, but exhibits a sustained cost saving after six months.

Other things equal, that pattern supports the 'encouraging prevention rather than cure' hypothesis on the effect of VHC.

Chart 8 – Allied health & GP costs initially increase, then trend down

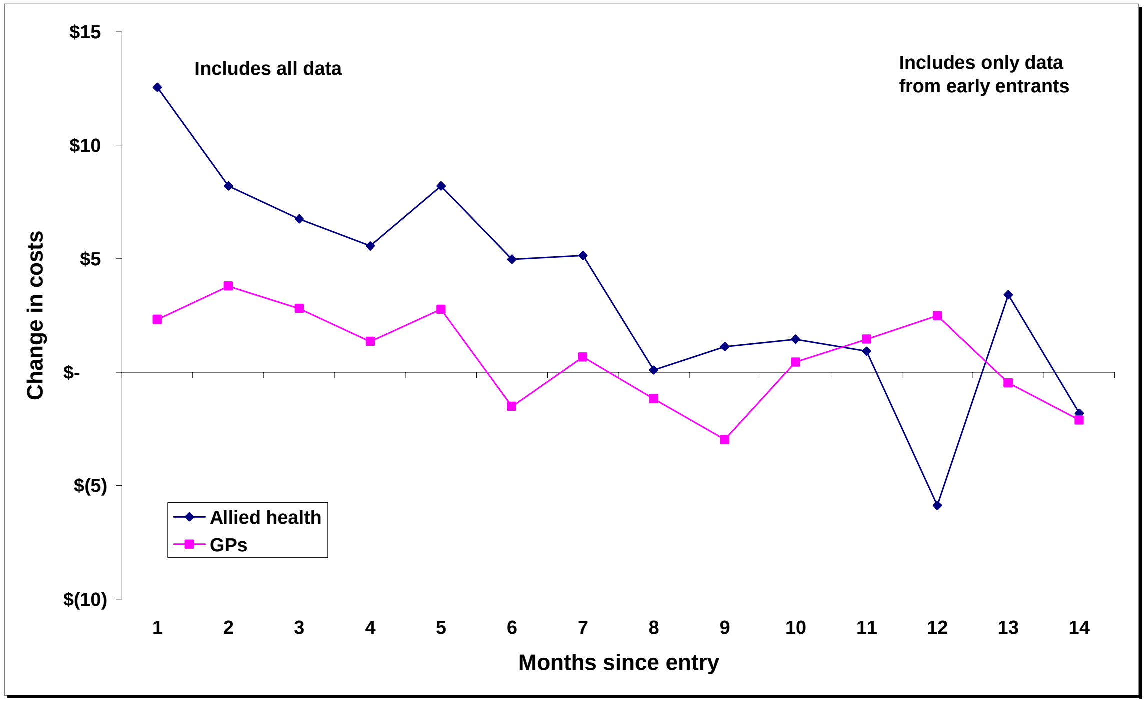
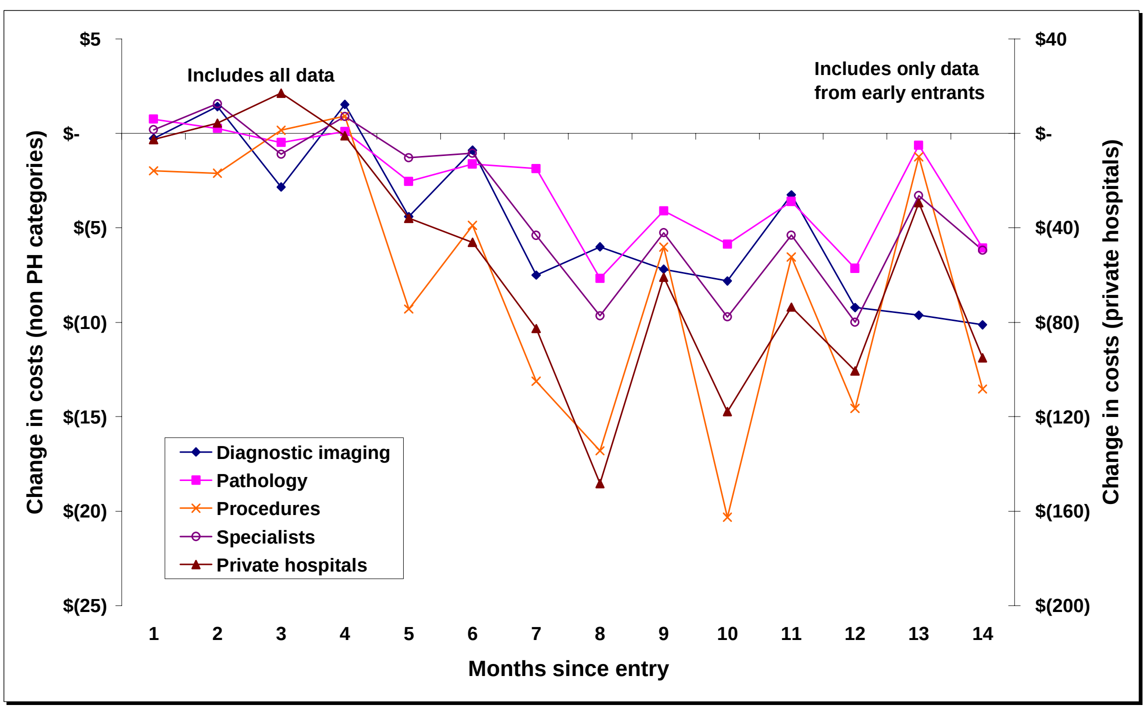


Chart 9 – Other categories of spending are initially unchanged, then they too fall



Later sections show a breakdown of these broad trends by both class of person (male / female, non-transitional / ex-HACC) and by month of entry into VHC. These later charts show a volatility which obscures the general trend in the above, simplified results.

7.1 Trend in total costs

Total costs are mostly driven by private hospital spending, but the early months record a cost increase due to allied health and GP spending. Overall, that points to a long term sustainable net saving.

Chart 10 – Downtrends over time also visible across categories of clients

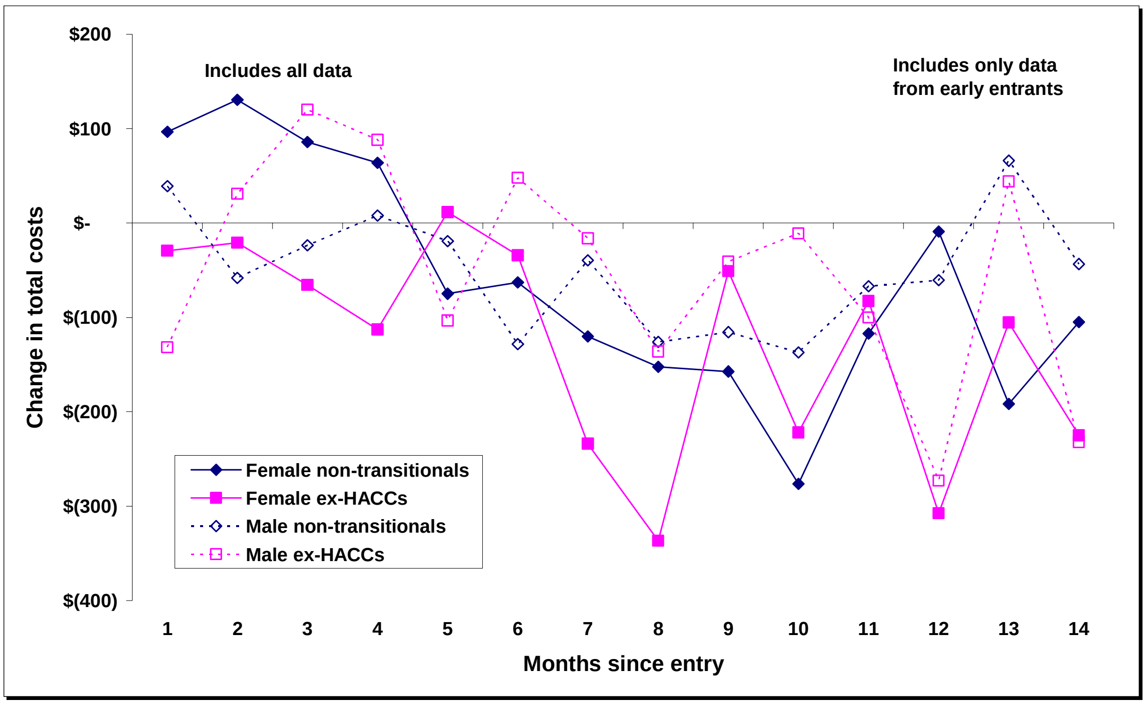
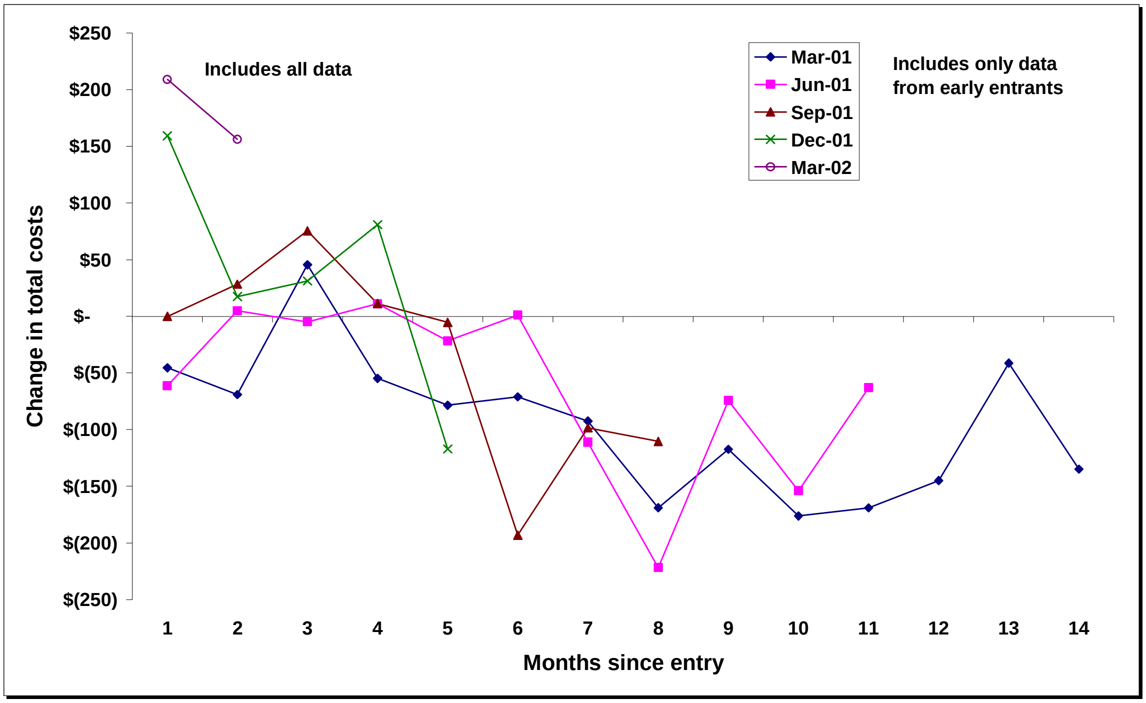


Chart 11 – With the downtrend over time evident across entry cohorts



7.2 Trend in allied health costs

The short term (6 month and less) response appears to be a cost increase, especially for non-transitionals (the diamonds on the first chart). In the longer term, the net effect on monthly allied health spending seems to be near zero.

Chart 12 – Most client categories see an initial increase in allied health spending

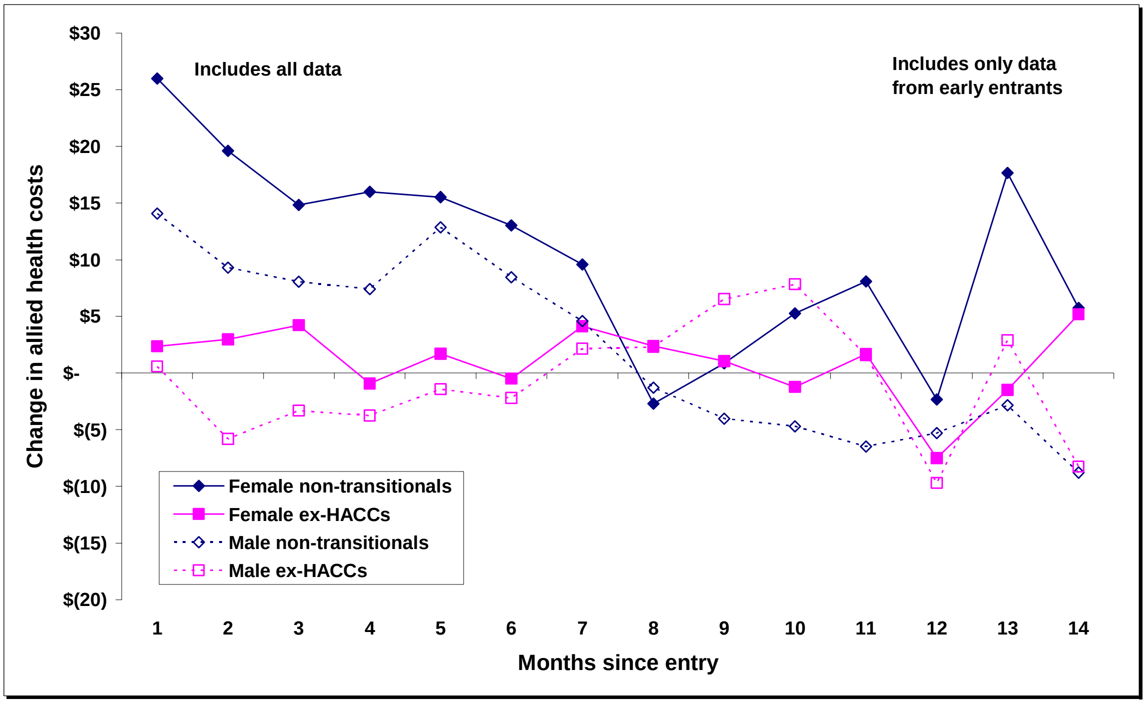
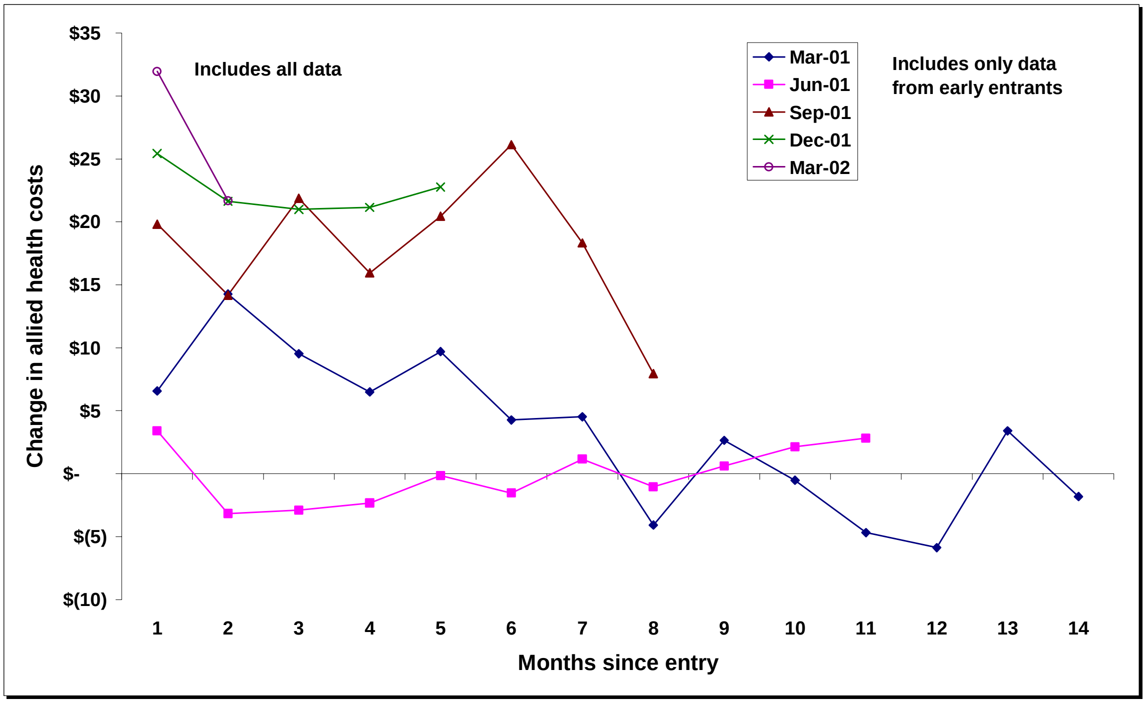


Chart 13 – The longer in VHC, the less the extra spending on allied health



7.3 TREND IN DIAGNOSTIC IMAGING COSTS

There is a general downward trend towards a cost saving, consistent across all four classes of person.

Chart 14 – Savings on diagnostic imaging increasingly evident across all client categories

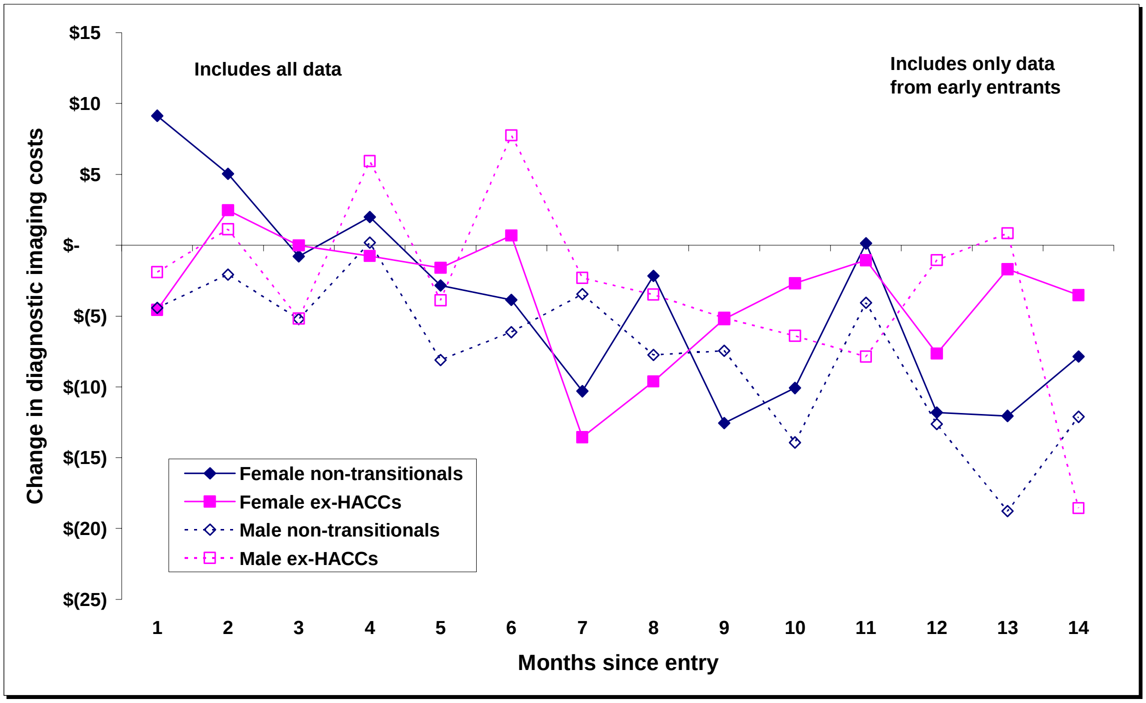
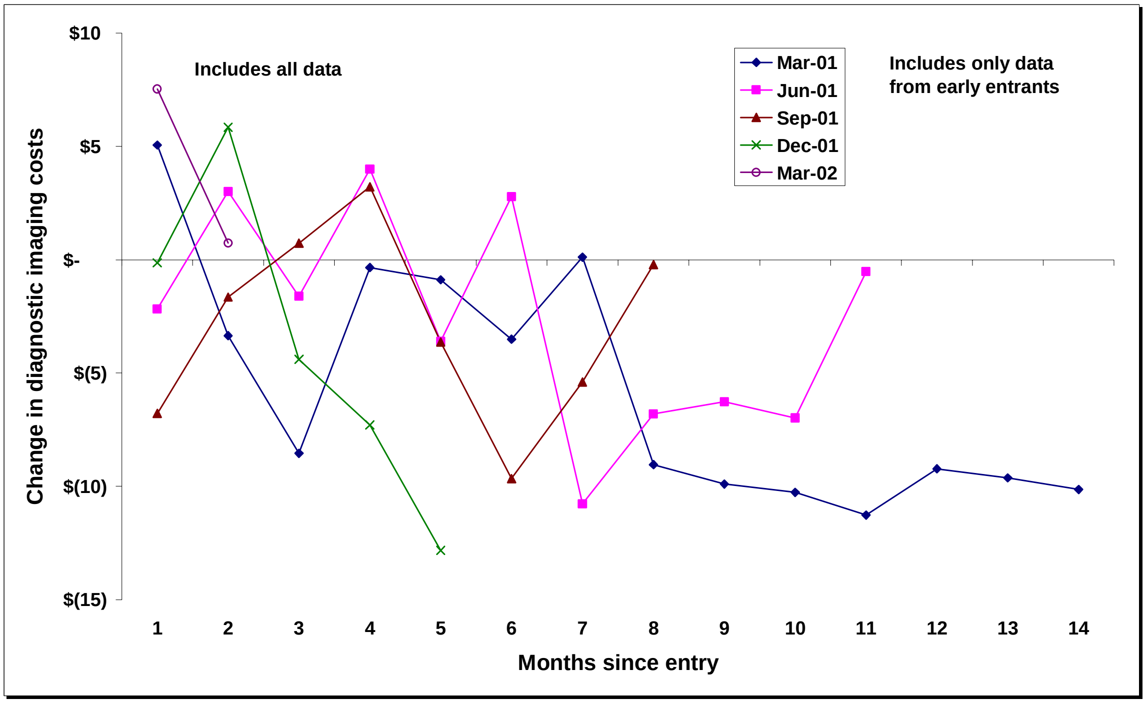


Chart 15 – The longer in VHC, the less the extra spending on diagnostic imaging



7.4 TREND IN GP COSTS

Similar to allied health spending, GP costs seem to rise in the short term, but move towards no change in the longer term. The variance in our estimates rises towards the end, as we have less data.

Chart 16 – Spending on GPS higher initially, little change over longer term

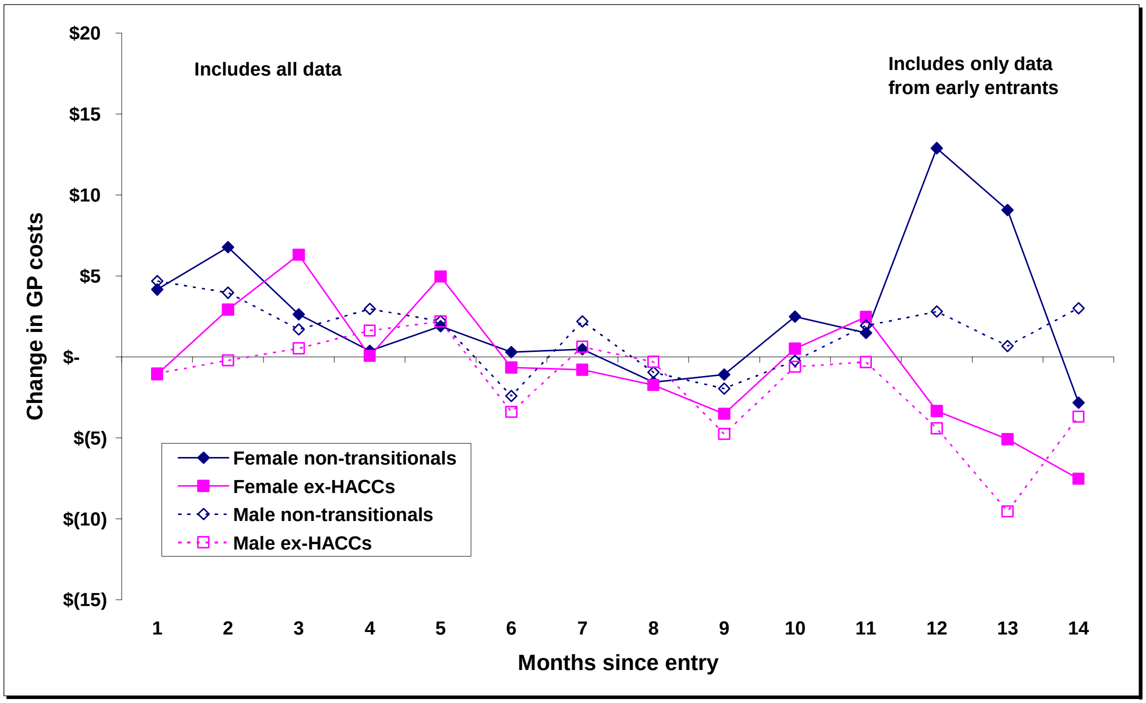
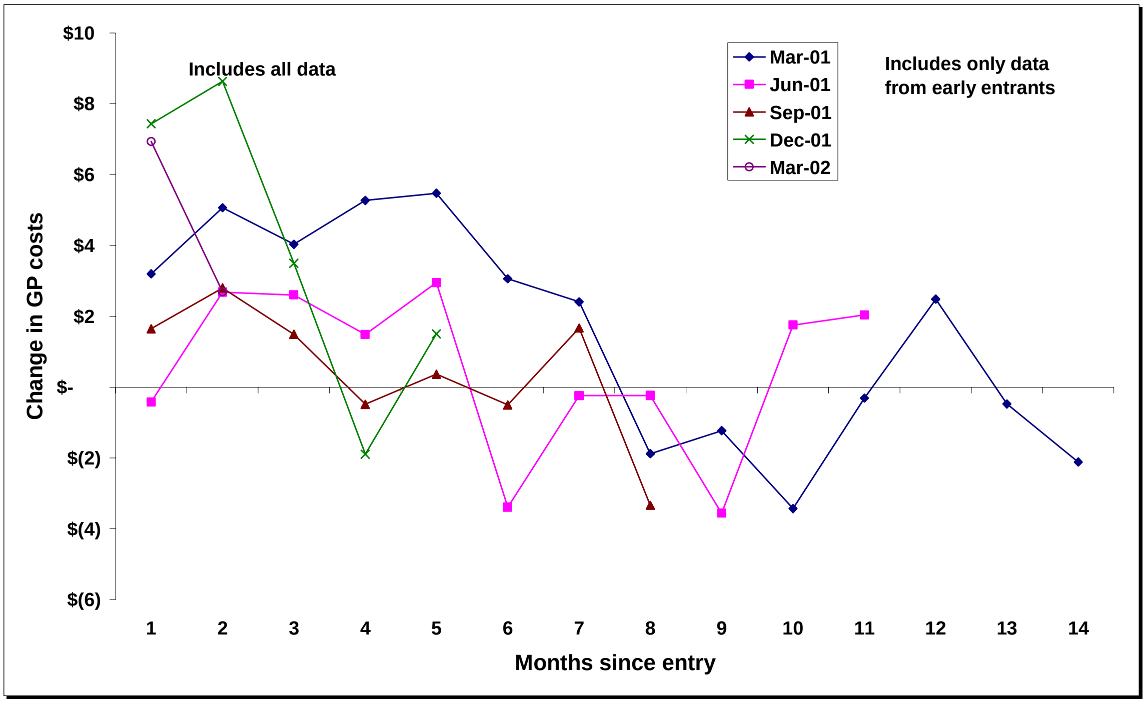


Chart 17 – GP spending trending down after initial rise on entry



7.5 Trend in pathology costs

There may be some sort of downward trend in pathology costs, but it seems to be quite volatile within a particular class (such as 'female non-transitionals').

Chart 18 – Volatility high among client category for pathology costs

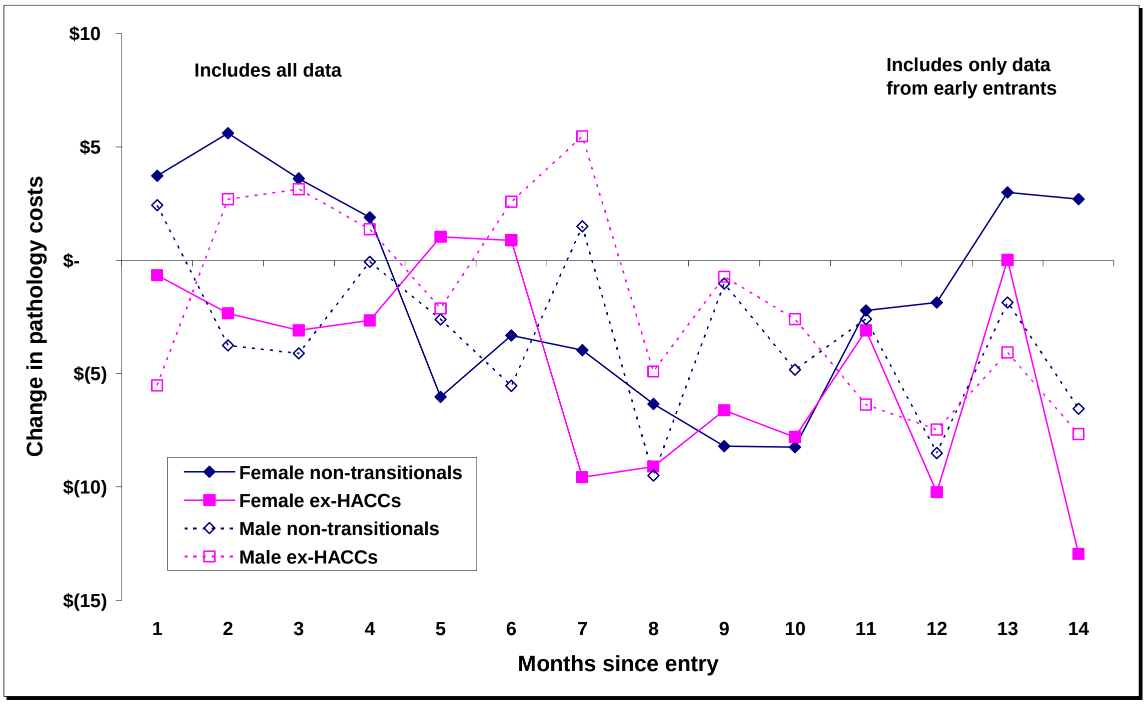
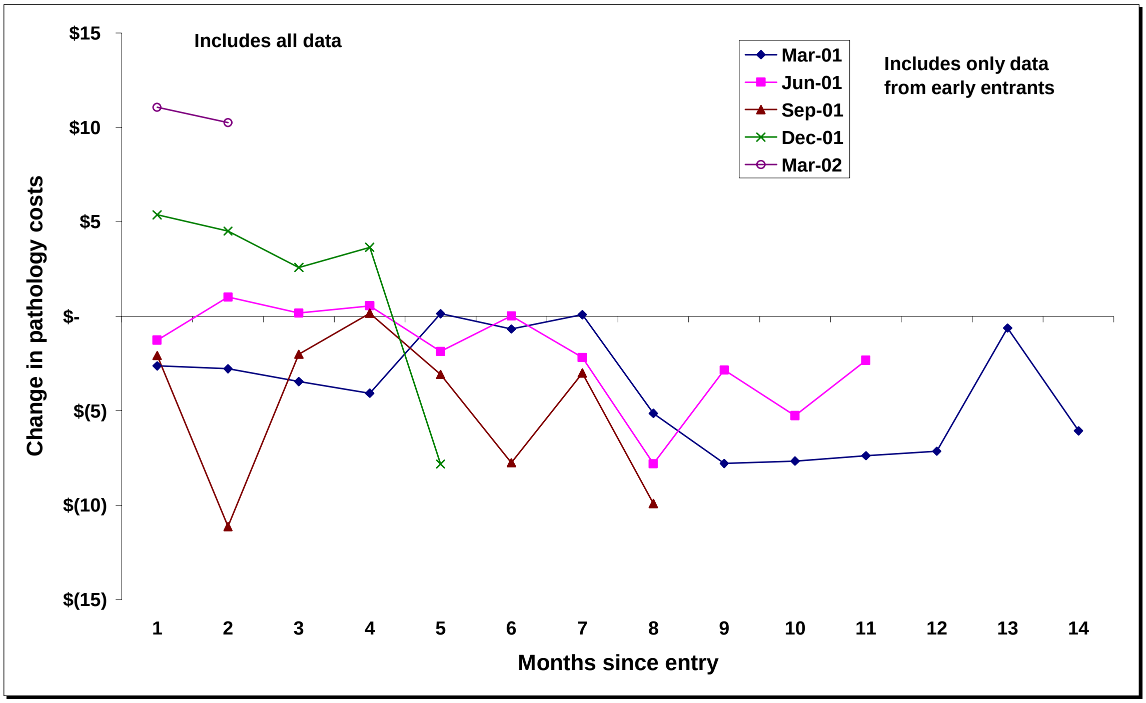


Chart 19 – Spending on pathology lower over time when examined by entry cohort



7.6 TREND IN PRIVATE HOSPITAL COSTS

Private hospital spending is quite volatile, as are estimated changes in private hospital spending. However, there seems to be a long-term sustained trend towards saving, stabilising at about the 6-months-after-joining-VHC mark.

Chart 20 – Spending on private hospitals trends down by client category, before settling at 6 months

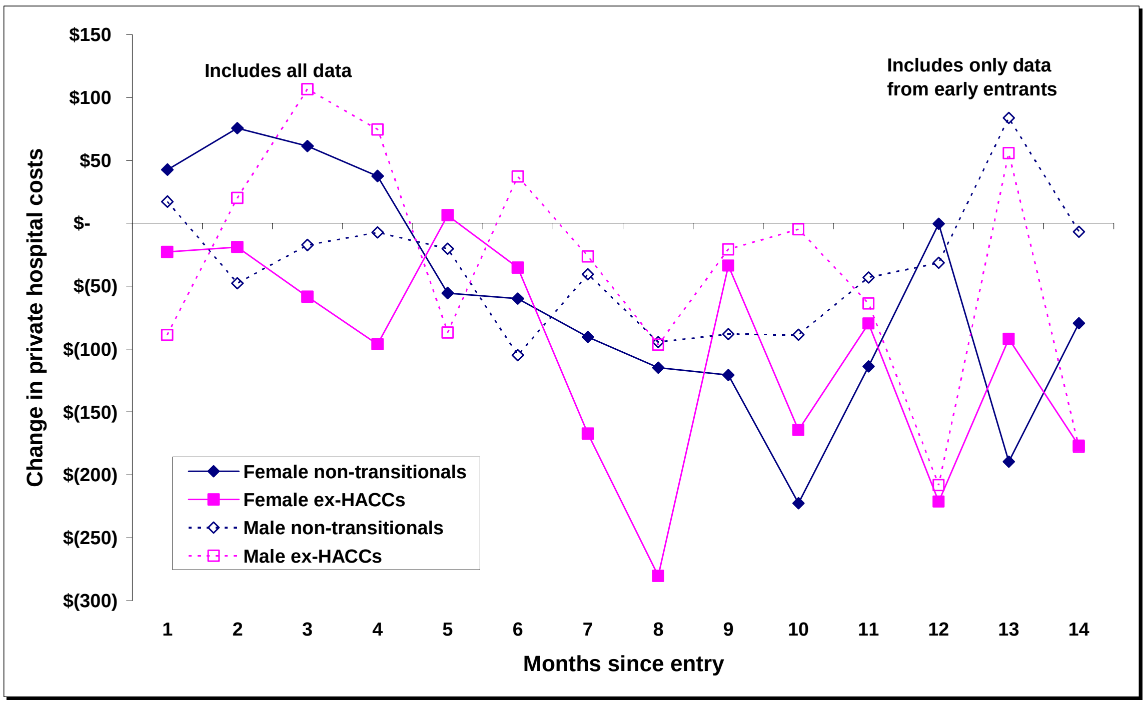
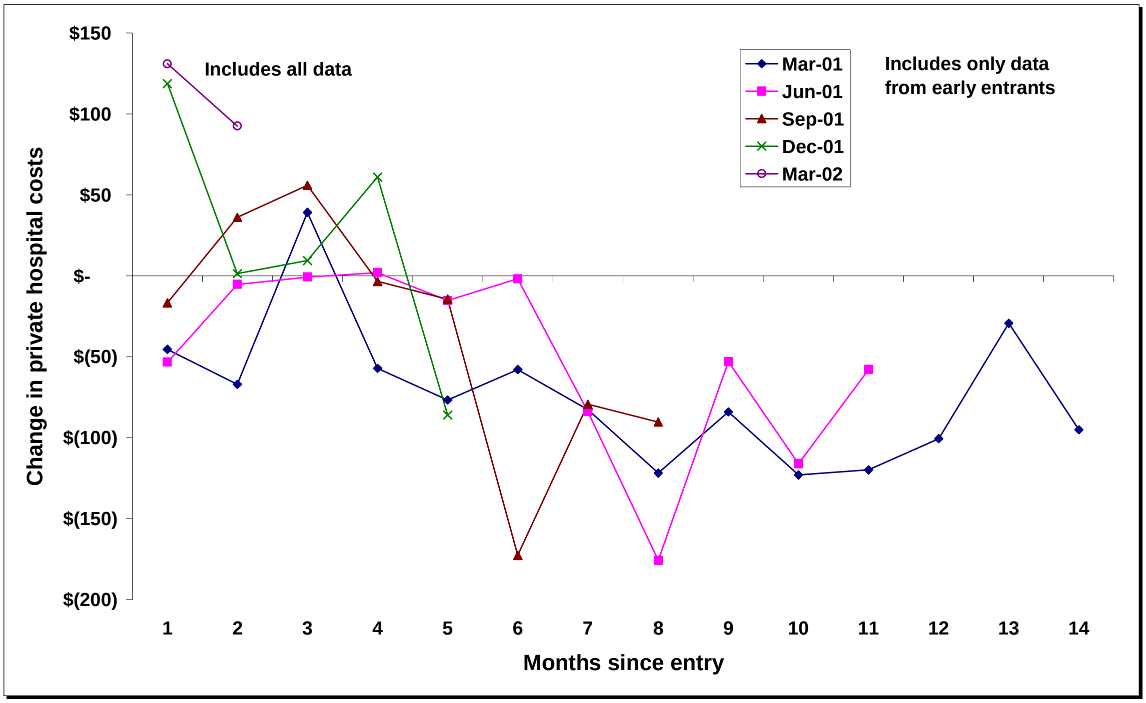


Chart 21 – Hospital downtrend now clear by entry cohort (after allowing for 'frailty')



7.7 Trend in procedures costs

Again, there is a downward trend, but there is quite a lot of noise, or variability, in the charts.

Chart 22 – Spending on procedures dominated by volatility in trend savings

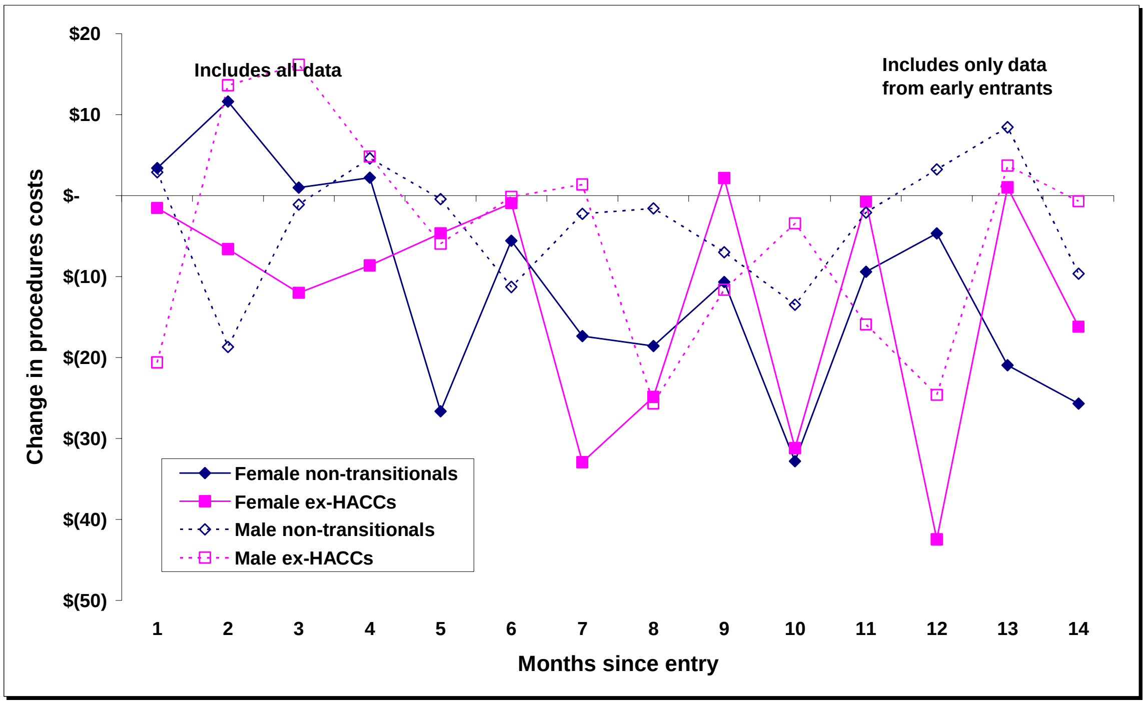
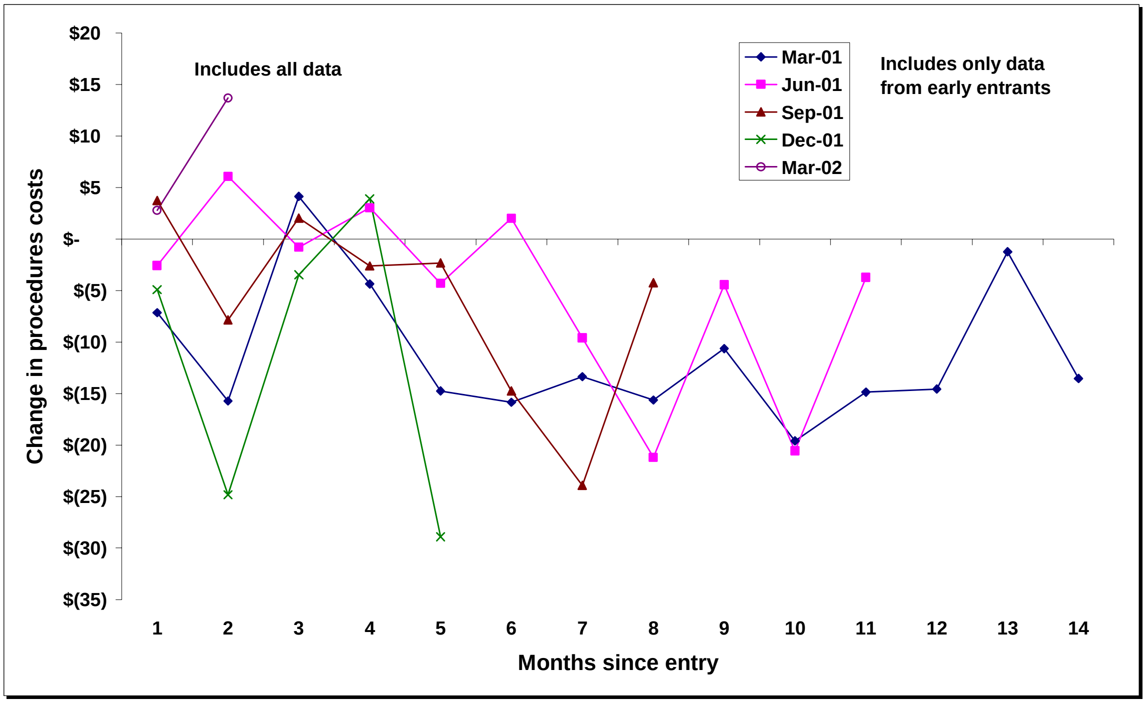


Chart 23 – Spending on procedures moves lower after entry



7.8 Trend in specialist costs

Similarly to private hospitals, things seem to settle to a cost saving at about the 6 month mark, although it is still quite volatile.

Chart 24 – Spending on specialists again displays steady downtrend by category

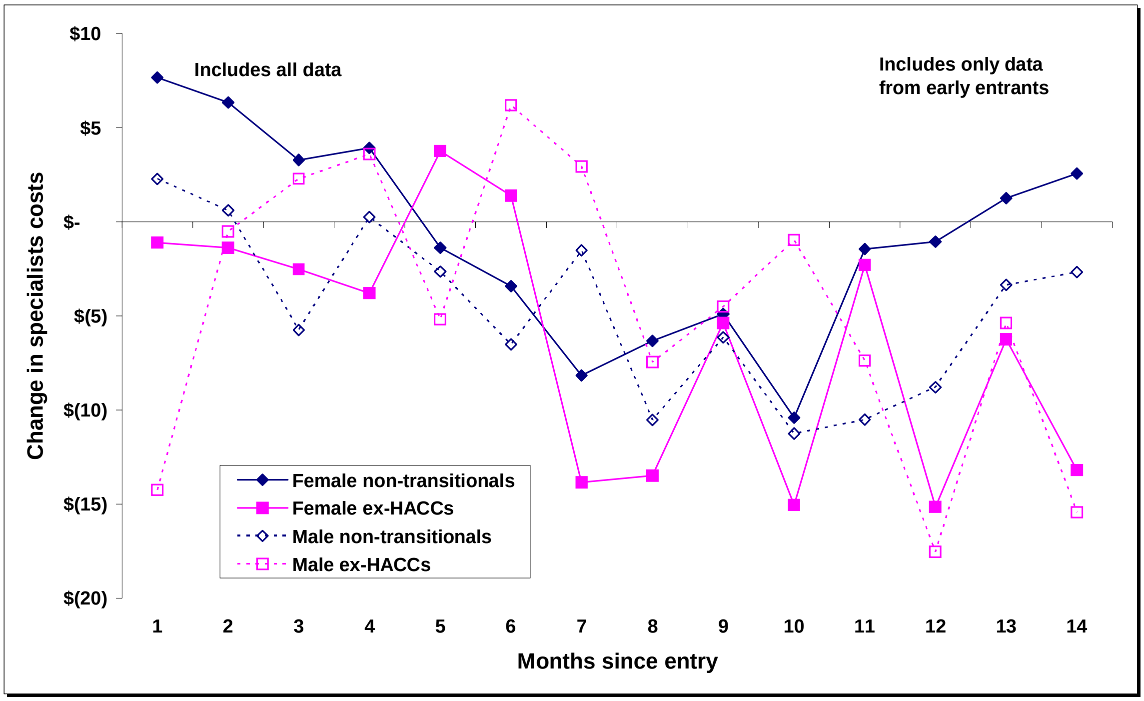
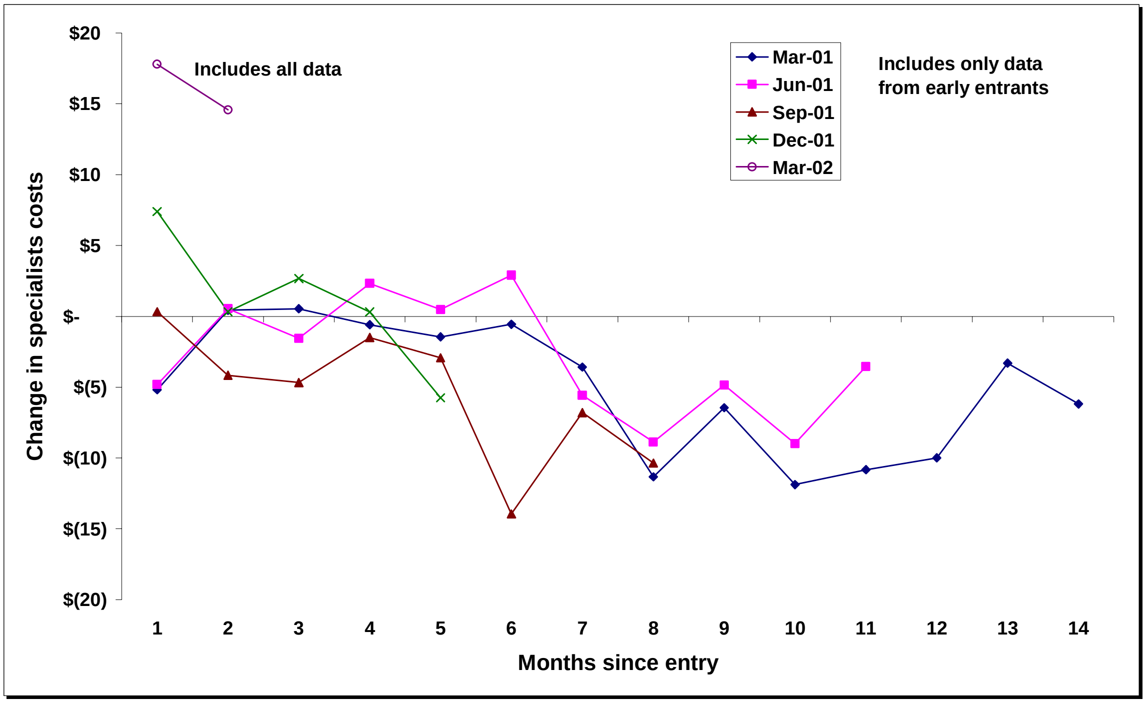


Chart 25 – Spending on specialists shows slow but steady downtrend by entry cohort



7.9 EXAMINING THE TRENDS

There is a clear pattern across time, with spending initially higher, but settling into a net cost saving after six months. As raised earlier, two possible explanations of this are:

1. Earlier cohorts (those people who joined the VHC programme earlier) dominate the long-term trend. These people may have been different in some significant way, such as being more or less frail (and thus affecting the comparison group they were matched with), or having been 'cherry-picked'. Most of the HACC transferees were moved into VHC in June and July of 2001, and in the months leading up to then, and that may affect the results.

2. We simply have more data for earlier cohorts – and that there is a short term cost increase but a long term cost saving (which is only apparent for earlier cohorts).

The first concern is that cohorts may be quite different, especially in terms of their 'frailty', or 'sickliness'. Our analysis incorporated individual disability pension rate data, as a proxy measure of 'frailty', and it appears that the distribution of disability variable (see table below) is fairly constant across cohorts. The slight pattern is that, in non-transitionals, the proportion of specially and extremely disabled entrants falls, and the proportion of non-disabled entrants rise.

Non-Transitionals	Mar-01	Jun-01	Sep-01	Dec-01	Mar-02
None	51%	54%	57%	59%	60%
Low (5%-95%)	20%	19%	19%	20%	20%
Medium (100%)	10%	11%	11%	10%	10%
Special	9%	8%	5%	5%	5%
Extreme	10%	8%	8%	7%	6%

Ex-HACCs	Mar-01	Jun-01	Sep-01	Dec-01	Mar-02
None	66%	63%	57%	61%	69%
Low (5%-95%)	15%	16%	14%	19%	15%
Medium (100%)	8%	9%	12%	7%	6%
Special	4%	4%	7%	5%	3%
Extreme	7%	8%	9%	7%	7%

Another point to note is that, in the earlier charts, there was not a consistent difference between different entry months (Mar-01, Jun-01, etc.), over the time period in which they are comparable. The exceptions are in allied health (where the five lines in Chart 13 do not seem to intertwine), and in pathology and specialists (where the spending on the Mar-02 intake is a little higher than the others in Chart 19 and Chart 25). Overall, though, the time trends seem to be consistent regardless of entry cohort.

8. INTERPRETING THE RESULTS

8.1 SUMMARY OF RESULTS

The results suggest that:

1) The VHC programme is associated with overall savings in health care spending.

2) Within the total, spending for non-transitionals in allied health and GPs is estimated to be higher than would otherwise have been the case.

3) However, spending in those categories is not higher for ex-HACCs. This dichotomy is one of the conclusions about which there is most confidence.

4) Spending for private hospitals and procedures is lower, with the change being concentrated in the female ex-HACCs. The volatility of private hospital spending means that, in most cases, the estimate of saving is not statistically significant, even though it is often large.

5) The most striking trend in relative time is towards larger savings the longer the time since the individuals entered into VHC. This result, evident once 'frailty' is allowed for, implies that there are larger estimated savings for earlier cohorts than for later ones.

6) The fact that the number of post-VHC entry months is less for later cohorts means that the relative time effects cannot be separated from the cohort effect. To do this will require longer post-VHC entry spending experience for the later cohorts.

7) At the level of the four classes, aggregated over spending categories and months, estimated standard errors across the classes are comparable. The statistical significance of the estimates largely depends on the relative sizes of the estimates of savings. Savings in private hospitals and the associated categories is only significant for female ex-HACCs.

8) Extreme observations are more likely to be from individuals with large savings.

Furthermore:

9) The various choices made in the course of the analysis have an impact on the estimates. An obvious example is the choice of the length of the base period. Decreasing the length of the base period gives larger estimates of savings. Mechanically, this largely comes from increasing some of the values of pre-VHC spending. When pre-VHC spending is subtracted from post-VHC spending to get the estimate of savings, increasing the former implies an increase in savings. But

it is also the case that the distribution of savings by individual has more outliers and is more skewed when the base period is shorter. These features of the distribution make its interpretation much more difficult. The impact of the choices is not factored into the standard errors.

10) Other caveats on the results have been noted.

8.2 INTERPRETATION

Parts of the August 2002 report focussed on a set of hypotheses for explaining the impacts of VHC. Three hypotheses, updated for this report are:

1. That the increase in spending is preventative – there should be small and on-going increases in 'preventative' categories of spending and subsequent larger decreases in spending in categories associated with 'treatment' or 'cure'.

2. That the VHC individuals are more sick or frail than non-VHC individuals with similar health spending histories.

3. That those in VHC receive more intensive servicing.

Which of these are still relevant?

8.3 PREVENTATIVE SPENDING

One hypothesis explaining the patterns of spending is that VHC is encouraging spending on 'prevention' rather than 'cure'. At the time of the August 2002 report, this hypothesis could not be rejected – the higher spending in categories associated with 'prevention' was in evidence, but it was too early to see spending savings in the other ('cure') categories.

After allowing for a proxy for 'frailty', there is now some additional evidence to support the hypothesis. First, non-transitionals show increases in spending on allied health and GPs, and those increases appear to be larger for later cohorts. The differences between non-transitionals and ex-HACCs are consistent with the hypothesis that ex-HACCs were already receiving this type of care. Second, private hospital spending appears to decrease over time.

Against this evidence are the facts that the higher level of spending in the preventative categories appears to trail-off over time and that the level of private hospital spending is volatile and the estimates are not statistically significantly different from zero (at the 5% level). The inability to separate the effect of time since entry into VHC from the cohort effect makes it impossible to conclude either way on this hypothesis.

8.4 Frailty

The second hypothesis is that VHC individuals are more frail or sick than non-VHC individuals with similar *observed* characteristics. The related question is: How much are the *unobserved* characteristics leading to entry into VHC and to subsequent spending taken into account in the methodology?

In this report we have set out to capture 'frailty' effects via the inclusion of the disability pension rate as a proxy for 'frailty'.

The underlying problem is that the available data only tells us so much about people. For people otherwise equivalent on the data we have (they are matched on age, gender, disability pension rate, and health spending prior to VHC entry), it may still be that those going into VHC are 'more frail' than those not entering. This is especially the case for females, since most have disability pension rate equal to zero.

However, health care spending is highly correlated with the disability variable that enters into the models. When the disability variable is added to the models, the estimated savings associated with VHC increases. These results are consistent with the disability variable being a useful proxy for at least some aspects of 'frailty'.

Of course, there may be other aspects of 'frailty' that remain unobserved. The methodology controls for aspects of 'frailty' that do not change over the time period of the analysis. But remaining aspects may bias the results towards an overstatement of the cost of VHC.

Recall that it is the process of selection into VHC, and possible correlation between that process and subsequent spending, that raises the possibility of sample selection bias. Our assessment of the possible estimation methods led us to the non-parametric, matched DID, estimator. This estimator is able to control for observed variables leading to the correlation and for unobserved variables leading to correlation – but only so long as the impact of the unobserved variables is constant over time.

However, temporary unobserved variables and shifts in the level of the unobserved variables may cause bias, if the variables are correlated with spending. Suppose there is a sudden change in the level of 'frailty' which quickly leads to entry into VHC. The change may not be reflected in spending data. By definition, the individual is able to stay at home, but he or she may require additional assistance from an allied health profession or a GP. The additional assistance could either be short term or long term, depending on whether the individual returns to normal.

The increase in spending would be incorrectly attributed to VHC. In the available data, the evidence of the change in the level of 'frailty' would be the higher spending. But that spending has already been attributed to VHC.

For those reasons there still may be aspects of 'frailty' affecting the results. In essence:

1) Our 'matching' process may be less than perfect, despite the sophisticated manner in which it has been conducted.
2) A sudden change in 'frailty' may not be reflected in the data. For example, for two otherwise equivalent people on the data we have up to the time of entry, the one who breaks a hip is the one likely to go into VHC. In the example just given, there would be a lingering increase in spending attributed to VHC, despite our careful matching of the 'before entry' experience.

Therefore bias in the estimate of VHC from the unobserved factors remains a possibility. But the likely bias is that the actual savings is greater than the estimate.

8.5 More intensive servicing

Of the statistically significant results, the clearest one is the increase in allied health and GP spending for non-transitionals. A possible hypothesis to explain that is 'for an equivalent set of symptoms or illness, those in the VHC may get more treatment than those outside'.

The data make it difficult to test for this effect and to distinguish it from the effects of 'frailty' – the higher spending in allied health and GPs could be from either more intensive serving or from 'frailty'. Essentially, we do not have data on the symptoms or illnesses which prompted treatment, nor on the processes by which individuals were assessed into VHC.

8.6 Summary

The composition of the data does not allow direct tests of the three possible explanations. The missing elements include:

- information on the interactions between individuals and the VHC and health care systems and information on the level of 'frailty' and illness of the individuals; and,

- a long enough span of data to be able to distinguish between possible early entrant effects and possible trends in spending over time

8.7 Overall assessment

The results here are dominated by the volatility of private hospital spending.

Addressing that volatility by having a six month base period for the comparisons between those in VHC and the 'matched' equivalent veterans outside it points to a saving of $15.10 per person per month.

Those results run the risk of missing the impact of missing variables – such as the thought that those selected to enter VHC are 'more frail' than those outside VHC, meaning that individuals in VHC are somehow more frail or sick than those outside of VHC.

Once an attempt to address the 'frailty' problem is allowed for – using disability pension rates as a proxy proved significant – the estimated saving rises to $21.32 per person per month.

The allowance for 'frailty' also saw a trend emerge in the results – one consistent with the hypothesis that VHC is seeing more 'preventative' spending early, and that extra investment is paying off via less 'treatment' or 'cure' costs later.

It would be risky to place too much weight on these results. As noted, they remain dominated by the volatility of private hospital spending. There is therefore a reasonable chance that both important results – the overall saving and the emerging evidence of successful 'preventative' spending – may simply be statistical noise.

Indeed, the most robust result is an increase in spending on allied health, rather than the overall fall in spending or the large fall in private hospital spending that dominates it.

That said, the overall results point to a saving to the Commonwealth Government from the VHC programme, and one that is growing over time.

9. SERVICE NUMBERS

We were also asked to consider changes in service numbers inside VHC. In brief, we found that changes in services numbers are generally small – less than one service per year per individual. These differences are not convincingly significant.

We have estimated the impact of VHC on service numbers for the categories of allied health, GPs, and specialists. Inspection of the raw data suggested that it would not be useful to estimate service numbers for the other categories. In those categories, there were often large numbers of services recorded as occurring on the same day, making it difficult to reliably estimate the true number of unique consultations.

The overall results are given in Table 9, which says that the estimated changes are all less than one service per three years. Only the change in GP services is somewhat statistically significant.

Table 9 - Overall changes in service numbers

Services / Year	**Estimate**	**Std Err**	**t-score**
AH servies	-0.3135	0.3279	-0.96
GP services	0.2874	0.1116	2.57
SC services	-0.1275	0.1368	-0.93
Observations	15539		

The breakdown of these numbers by class of person is presented in Table 10. These results show no consistent patterns – for example, the estimated change in allied health services are an increase for female non-transitionals and male ex-HACCs, but a decrease for female ex-HACCs and male non-transitionals.

A number of individual estimates are statistically significant in isolation (e.g., specialist services for female ex-HACCs), but there is no agreement across any one class of person, or across any one category of spending. Overall, these detailed results point to no convincing conclusion about the direction or magnitude of the effect of the VHC programme on service numbers.

Table 10 - Changes in service numbers by class of person

	Female non-transitionals			Female ex-HACCs		
Services / Year	**Estimate**	**Std Err**	**t-score**	**Estimate**	**Std Err**	**t-score**
AH services	0.6406	0.6653	0.96	-2.1381	0.4815	-4.44
GP services	0.6722	0.2184	3.08	-0.0651	0.2089	-0.31
SC services	0.6042	0.2693	2.24	-0.9853	0.2342	-4.21
Observations	4329			3329		

	Male non-transitionals			Male ex-HACCs		
Services / Year	**Estimate**	**Std Err**	**t-score**	**Estimate**	**Std Err**	**t-score**
AH services	-0.4906	0.6621	-0.74	0.7577	0.6683	1.13
GP services	0.4568	0.2138	2.14	-0.2262	0.2342	-0.97
SC services	-0.1390	0.2625	-0.53	-0.2218	0.3087	-0.72
Observations	5206			2675		

These changes in service numbers are generally consistent in sign with the change in actual spending as well. For example, specialist services are estimated to rise, for female non-transitionals, but to fall for all other classes. This pattern repeats in specialist spending. Similarly, both the estimated impact on GP service numbers and on GP spending is higher for non-transitionals.

10. ON FORECASTING

An important question is whether the results in this report are able to support forecasts of future trends in VHC. The essential point to note is that the model used here is an evaluation model, rather than a forecasting model. The key question addressed in this report is whether spending is higher or lower than it might have been in the absence of VHC.

The estimation of this impact uses a matched DID estimator, which we judge to be the most suitable for the task. The estimator compares changes in spending for VHC individuals with changes for similar non-VHC individuals.

The key advantage of modelling of changes rather than levels means that there is no need to consider factors that have a constant effect on spending through time. Similarly, factors that imply trends in spending, but equal trends for VHC individuals and non-VHC individuals, can be ignored. Indeed, being able to ignore these factors is one of the main advantages of the DID estimator. The DID estimator also ignores non-VHC individuals who are not matched (apart from in the estimation of the propensity model).

The advantages of this approach therefore lie in answering the question of whether spending has risen or fallen. However, the resultant focus on changes rather than levels means that the analysis in this report says nothing about the underlying trends in spending. That therefore limits its usefulness in forecasting.

But what can be said about future impact of VHC? The analysis suggests that:

- the observed characteristics of the individuals entering VHC are fairly stable over time; and
- the propensity models are not stable over time; but,
- the results are consistent with moderate, persistent, and fairly precise, estimates of differences in spending in VHC in some categories (such as allied health); and large, but variable and imprecise, estimates of differences in spending in some other categories (such as private hospitals). The latter categories also dominate total spending.

It seems reasonable to expect that such features will continue, which implies that continuing volatility in private hospital spending will ensure that it remains difficult to predict the impact on spending of subsequent experience under VHC (relative to what might otherwise have been).

What can be said about forecasting overall spending? We noted earlier that within aggregate spending there is a distinct pattern: the spending of VHC individuals are increasing on average whereas those of non-VHC individuals are relatively flat. These relative movements in themselves say nothing about the trend in overall spending, just how the overall movements are split between the two parts.

That implies that a careful preparation of separate forecasts for VHC and non-VHC would be needed in the context of those relative movements. For example, non-VHC may be flat because VHC is rising. It might also take into account the sample selection, the small increases in some categories, and the large variances in others. To the extent that the relative movements cancel out, it may be better to ignore the split and forecast total spending.

But taking all factors into account in forecasting overall spending is clearly a major exercise. These factors might include the number of individuals entering the population of card holders, the number of deaths, any rule changes, any changes in the behaviour of health service providers in the face of different rates inflation of costs and DVA reimbursements, etc.

Of course, a rough forecast of overall spending could be obtained by simply extrapolating past trends. The greater complexity of the model would aim at the relatively small, hard-to-get increases in accuracy.

VHC also had some impact on spending. As we have argued, it may be difficult to accurately estimate the effect of VHC on the macro data. A simple way to estimate the effect would be from forecasts of spending in the VHC period based on pre-VHC data. The estimate of the effect is the difference between actual and forecast spending. But:

- there are relatively few pre-VHC observations from which to model trends – about three years of data; and

- a small change in the trend could lead to a large change in the estimate of the impact.

We suggested elsewhere that, because VHC has been operating for a relatively short time period, the programme may not have yet settled into a steady-state. But some of the evidence in this report suggests otherwise. For example, the observed characteristics of individuals are not changing significantly.

Additional questions that might be considered in the preparation of macro forecasts include:

- Should the total be forecast or should the data be split into classes of individuals (males, females, ex-HACC, etc), with the forecasts of the individual classes aggregated into the total?

- Should the age of individuals be factored into the forecasts?
- Should regions or State/Territories be factored into the forecasts?

In summary, the evaluation model in this report is not well suited to the task of forecasting trends in VHC. Furthermore, the questions raised in this section suggest that obtaining accurate forecasts of overall spending in VHC, especially one useful for estimating the historical impact of VHC, would be a large exercise.

11. FINAL THOUGHTS

Although our comments have not been called for beyond an evaluation process, we take this opportunity to offer some final thoughts.

There are a number of commendable design features of VHC. The purchaser/provider split is good practice. It provides clear lines of accountability, is conducive to proper internal control, and lends itself to effective resource allocation and monitoring.

These arrangements exert significant positive accountability pressures but there may be areas for potential improvement. In the main these go to getting a better handle on the pressures over time for additional spending and assessing whether short term preventative spending would produce long term returns to the Budget. These issues are discussed below.

11.1 CONDUCTING CONTROLLED STUDIES

There are a number of benefits in terms of conducting additional work using longitudinal studies to determine the savings versus the costs of early programme intervention and from different types of spending.

For instance, the findings of Anna Howe in her report *Targeting in the Home and Community Care Programme*, Aged and Community Care Services Development and Evaluation Reports, No 37) for the National Ageing Research Institute and Bundoora Extended Care Centre (1999), found significant differences in health outcomes for a marginal increase in the level of home help in HACC. The report found that after 18 months of being observed that 72% of the control group remained at home compared to 86% of the intervention group, with the control group having worse outcomes on several measures:

- more had died (11% compared to 4%);
- more were admitted to residential care, either nursing home or hostel (10% compared to 6%);
- more could not be traced, suggesting a change in living arrangements (7% compared to 2%);
- more reported major new health problems (45% compared to 32%);
- more had been hospitalised (47% compared to 35%) and were more likely to have had hospital stays of more than six days (36% compared to 29%); and
- more reported worsening dependency in relation to need for heavy housework (8% compared to 5%).

Developing devices to systematically track such changes in health outcomes for a change in key inputs would, over the long term, help the ongoing assessment of VHC outcomes.

11.2 Potential HACC transferees

Access Economics believes that DVA and the Department of Health and Ageing may benefit from the investment of more resources to determine the number of veterans still to transfer to VHC and the likelihood and potential rate of them doing so.

11.3 Assessment processes

Given the intensive nature of VHC, it seems reasonable to conjecture that assessors and providers tend to develop close relationships with 'their' veterans and war widows. This comes from the tight assessor/provider market but, more importantly, effective in-home caring often requires such a close relationship to be built and maintained with the individual. The pre-eminence of the veterans and war widows is reinforced with guidelines for the programme that make it clear that client interests are paramount.

This convergence of environmental influences may have the potential to expose the programme to some risks. Good quality monitoring at DVA Head Office (with on-the-ground input from State offices) is important in identifying any 'problem' assessors and to check the quality of service delivered by contracted service providers.

In the first round of analysis, DVA indicated that about 5% of veterans assessed for the first time for VHC services were not accepted into the programme. We understand that more detailed data analysis across the 54 VHC regions for the past 12 months shows that this rate ranges from 0-14%. Some unit cost analysis has also been undertaken. Region-by-region results have been advised to DVA State offices for follow up with particular assessors. The provision of such information, including hours by service and trends, may well be useful tools for State offices in their discussions with assessors to promote consistency of practice across regions. There may also be scope for further analysis. This would cover:

- Defining characteristics of veteran and war widow populations. Might local compositional differences go some to explaining differences in unit costs between regions?
- How are referrals occurring, and what are the differences in referral rates? In the context of this report, information on referrals and the method could be a key ingredient in the matching method.
- Examining the use of the Standard Assessment Instrument and comparing rates of acceptance into the programme when it is not used.

- Assessing the circumstances in which service upgrades occurring (regular, extraordinary, and who/what are the trigger agents)?
- What are the regional differences in the referral rates out of VHC into other non-DVA programmes that may better suit the particular requirements of the veteran and war widow?
- How many veterans and war widows leave VHC for other programmes and for Residential Aged Care facilities each year? What does an analysis of the regions show?
- Among veterans and war widows in receipt of VHC, what is the extent of service upgrades?

11.4 Monitoring service caps

The separation of the assessor and the service provider functions under VHC provides natural protection against over-servicing. There are very few services where this split does not exist. With this structure and with the fixed fee for service arrangements, the scope for 'ratcheting up' service levels over time is relatively limited.

The VHC Guidelines also put the emphasis on ensuring that, before a service level upgrade occurs, consideration should be given to whether the veteran or war widow would be more appropriately accommodated in another DVA programme, such as community nursing, or in another Commonwealth programme, such as a Community Aged Care Package (CACP).

Higher levels of programme spending that show up in particular regions may be an indicator that the veteran may be inappropriately housed in a DVA programme. But this could just as easily point to blockages in other parts of the system. For example, trouble accessing a CACP could cause a veteran to remain (inappropriately) in VHC. That, if combined with an attitude of keeping the veteran or war widow at home at all costs, may be against the long-term interests of the veteran or war widow and taxpayers.

Access Economics
21 January 2003

TECHNICAL APPENDIX

This technical appendix details the econometric methods used in the analysis of the impact of VHC on other programmes of the Department of Veterans' Affairs. The impact is measured through the spending on individuals in a set of seven health care spending categories – allied health, diagnostic imaging, general practitioners, specialists, pathology, private hospitals, and procedures.

The VHC programme began in January 2001. This analysis uses data up to and including May 2002. We focus on particular months – March 2001, June 2001, September 2001, December 2001 and March 2002.

We refer to these months as the 'decision months'. Thus, for individuals entering VHC in March 2001, there are fourteen months of post-entry data. For individuals entering in March 2002, there are only two months of post-entry data. The sample of months is used to speed the computations, while still covering the full period in which VHC has been operating and complete data is available.

Spending data is available at the individual level, by very detailed service type. Services are aggregated into the seven categories. Our approach, based on the individual records, makes direct use of this rich dataset.

The estimates of the impact of VHC are based on comparing the levels of spending before and after the individuals enter into VHC. However, it is not sufficient to simply estimate the average increase (decrease) in spending. This is because there may be correlations between the process that determines entry into VHC and spending patterns. For example, suppose that entry into VHC is more prone to happen following a temporary increase in spending. Then an associated decrease in spending is likely to occur after entry and it would be incorrect to attribute that decrease to VHC. More generally, the analysis of spending should take into account the process of entry into VHC and correlations between entry and spending.

For individuals who enter VHC, the month of entry is taken to be the month in which care commences. Information on the home care the individuals are assessed to receive is available in the dataset (the 'hours' variable), although we do not make direct use of this information in the econometric analysis. This is because the econometric methods are based on comparing VHC and non-VHC individuals and information of this type is not available for the non-VHC individuals.

However, we note that some of the individuals are assessed to receive only a few hours of home care. The direct impact of a couple of hours of home care may be small, and presumably less than the impact from a large number of hours. Our modelling is based on the assumption that the impact of VHC is the same across individuals, so that we implicitly assume that, in some cases, it is the interaction with the VHC systems that leads to the impact on other programmes, rather than the care itself.

VHC provides home care, hence individuals not living at home are not eligible. However, individuals in residential care, etc., are included in the dataset, as non-VHC individuals, and there is no information to flag them. Hence, they are included in the analysis. This will bias the conclusions if such individuals have different spending behaviour from individuals living at home. According to information from DVA, there are about 18,000 veterans and war widows living in residential aged care facilities.

The outline of this appendix is as follows:

- Section 1 covers the specification of the model and the measure of the impact of VHC – 'treatment on the treated'.
- Section 2 presents the estimation and testing methodology.

1. MODEL SPECIFICATION AND TREATMENT ON THE TREATED

1.1 BASIC MODEL

We allow the impact of VHC to vary between males and females and between ex-HACCs and non-transitionals. Since gold card holders typically have higher spending than white card holders, we assume that the impact varies across card types. We concentrate on gold card holders. We assume that the impact does not vary with age.

For a particular type of individual and a particular category of spending, we denote the outcome variable (i.e., spending) for individual i in period t by Y_{it}. This outcome is assumed to depend on a set of exogenous variables, X; a dummy variable, d, signifying entry into VHC; and unobserved factors, U.

Focusing on month k, we define $d_i = 1$ if individual i enters VHC in month k and $d_i = 0$ otherwise. k represents the decision month. We assume,

(1) $Y_{it} = X_{it}\,\beta + d_i\,\alpha + U_{it}$ if $t > k$

(2) $Y_{it} = X_{it}\,\beta + U_{it}$ if $t \leq k$

In (1) and (2), we have assumed that the model is linear in the variables and that the parameters are constant over time and values of d. Setting $d_i = 0$ and $d_i = 1$ in (1) shows that α measures the impact of entry into VHC for individual i: after period k, spending for individuals in VHC ($d_i = 1$) systematically differs from that of non-VHC individuals ($d_i = 0$) by the amount α.

The set of parameters β define the relationship between the exogenous variables X and the dependent variable Y, and the error term U has mean zero.

The impact of VHC may vary over time, so we could write

$$Y_{it} = X_{it}\,\beta + d_i\,\alpha_t + U_{it} \qquad \text{if } t > k$$
$$Y_{it} = X_{it}\,\beta + U_{it} \qquad \text{if } t \leq k$$

Similarly, the impact could depend on the month in which a person enters VHC, if persons entering in different months are systematically different (although it appears from the results that this is not the case).

The X variables include age, health spending history, State/Territory of residence, and disability pension rate. The split by gender, card type, and prior status also implies that these variables are allowed for as the explanatory variables. Spending from the two months immediately prior to the decision month, but not before that, are included in X.

Entry into VHC is most probably not random. As discussed in our earlier report, this may lead to non-zero correlation between entry into VHC and the unobservables in the spending equation.

We assume that entry into VHC can be modelled as follows. For each individual, there is an index IN such that the higher the value of IN, the more likely is the individual to enter into VHC. IN depends on a set of exogenous variables Z:

(3) $$IN_{ik} = Z_{ik}\gamma + V_{ik}$$

and entry occurs if IN is greater than zero,

$$d_i = 1 \text{ if } IN_{ik} > 0$$
$$d_i = 0 \text{ if } IN_{ik} \leq 0$$

The value of zero is simply a normalisation.

The non-zero correlation between U and V is the cause of the sample selection effect, and may lead to bias in the estimation of the impact of VHC if it is not taken into account.

1.2 Individual effects

The available data is in the form of a panel, with observations on the set of individuals through time. It is common in such data to assume the presence of individual effects – individual specific intercepts (e.g., Wooldridge, 2002, Chapter 10). These effects represent variables/characteristics which are particular to the individuals and which do not vary through time. They are not observed by DVA (i.e., cannot be measured by variables in the dataset), and so are included in the error term. An example might be the degrees of 'frailty' of the individuals (though we attempt to provide a proxy for the latter through the disability pension rate).

We assume that U_{it} can be decomposed into

(4) $$U_{it} = \varphi_i + \mu_{it}$$

where φ_i is the individual effect and μ_{it} is a temporary individual-specific effect.

A similar decomposition can apply to V_{ik} in the selection equation (3). A possible assumption is that the correlation between the errors U and V comes from the individual effects, and that the temporary individual-specific effects are not correlated.

1.3 On estimating the impact of VHC

In the above model, the aim is to estimate α. A simple comparison of average spending for individuals in VHC and individuals not in VHC gives, in some month t after the decision month,

$$(5) \qquad \hat{\alpha}_t = \overline{Y}_t^{(1)} - \overline{Y}_t^{(0)},$$

where $\overline{Y}_t^{(1)}$ is average spending for individuals in VHC and $\overline{Y}_t^{(0)}$ is average spending for individuals not in VHC. $\hat{\alpha}_t$ estimates,

$$\begin{aligned}(6) \qquad & E[Y_{it}|d_i = 1] - E[Y_{it}|d_i = 0] \\ & = E[X_{it}\beta + d_i\,\alpha_t + U_{it}\,|d_i = 1] - E[X_{it}\beta + d_i\,\alpha_t + U_{it}\,|d_i = 0] \\ & = E[X_{it}|d_i = 1]\beta + \alpha + E[U_{it}\,|d_i = 1] - E[X_{it}|d_i = 0]\beta - E[U_{it}\,|\,d_i = 0]\end{aligned}$$

Thus, two effects are not controlled for:

- Any differences in the X's for VHC and non-VHC individuals.
- The sample selection bias: $E[U_{it}\,|\,d_i = 1] - E[U_{it}\,|\,d_i = 0]$.

The non-experimental estimators used in this analysis attempt to control for these two factors.

1.4 Censored data and other non-linearity

A number of forms of non-linearity could be present in the models. For example:

a) It was noted in the earlier report that it is possible to have zero spending in a category. In this case, the linear models in (1), (2) should be replaced by Tobit models such as censored regression models or Type 2 Tobit models (e.g., Amemiya 1985, Chapter 10). The Tobit models can be written as non-linear regression models, in which case, (1), (2) are replaced by non-linear forms. More generally, the linearity assumption in (1), (2) may be inappropriate.

b) The β parameters in (2) may be different from those in (1) if entry into VHC impacts on the sensitivity of spending to the explanatory variables.

c) In some of the spending categories, particularly private hospitals, there are a small number of individuals with very high spending – there are outliers in the data. As we shall see, these have importance consequences for the analysis. A possible modelling strategy involves taking natural logarithms of the spending data.

Allowing for these effects suggests the more general model,

$$(7) \qquad Y_{it}^{T} = g_{it}^{T}(X_{it}) + U_{it}^{T}$$
$$(8) \qquad Y_{it}^{C} = g_{it}^{C}(X_{it}) + U_{it}^{C}$$

where Y_{it}^{T} and Y_{it}^{C} are the outcomes for the VHC and non-VHC groups, which are written as general functions of a set of X variables plus unobservable error terms, U_{it}^{T} and U_{it}^{C}. (In (7) and (8), 'T' is for treatment and 'C' is for comparison.)

1.5 Treatment on the treated

What can be estimated – α, or some combination of α and other terms – is directly related to the assumptions made about the model. In particular, in a fully ***parametric*** model, in which functional forms are assumed for the $g_{it}^{T}(.)$ and $g_{it}^{C}(.)$ functions and a joint distribution is assumed for $\{U_{it}^{T}, U_{it}^{C}, V_{it}\}$, a variety of 'treatment effects' can be estimated. But in our earlier report we expressed concerned about some of these assumptions. In this report we also consider ***non-parametric*** methods – methods which avoid specific assumptions about functional forms and distributions. An implication of this is that fewer types of treatment effect can be identified. One effect that can be estimated is the 'treatment on the treated', defined as

$$(9) \qquad \alpha_{T} = E[Y_{it}^{T} - Y_{it}^{C} \mid d_i = 1].$$

Treatment on the treated represents the impact of VHC on individuals who were actually selected to VHC. The idea is to consider individuals who were selected into VHC, and estimate what their spending would have been had they not actually received any home care.

Of course, Y_{it}^{C} is not observed for individuals with $d_i = 1$. Selected observations with $d_i = 0$ are used instead. Combining (6) and (9) gives,

$$\begin{aligned}(10) \qquad & \{E[Y_{it}^{T} \mid d_i = 1] - E[Y_{it}^{C} \mid d_i = 0]\} \\ & = \{E[Y_{it}^{T} - Y_{it}^{C} \mid d_i = 1]\} + \{E[Y_{it}^{C} \mid d_i = 1] - E[Y_{it}^{C} \mid d_i = 0]\} \\ & = \alpha_{T} + \{E[Y_{it}^{C} \mid d_i = 1] - E[Y_{it}^{C} \mid d_i = 0]\},\end{aligned}$$

where the right-hand-side term in {} is the potential bias. Under specific assumptions, to be discussed later, the bias is zero.

Also, the treatment on the treated parameter is based on differences in averages – the non-parametric methods are also based on averages.

2. METHODOLOGY

2.1 ESTIMATION METHODS

A number of methods are available for evaluating the impact of a programme such as VHC. Here, we mention four of the methods. More detailed descriptions follow.

The methods are:

- *The Heckman Selection Estimator*
  - Based on a parametric model – requires assumptions on the functional forms of the model.
  - Gives the best approach if the functional form assumptions are correct.
- *The Matching Estimator*
  - A non-parametric method.
  - Matches each VHC individual with one or more similar non-VHC individuals, and compares their spending.
  - Susceptible to sample selection bias if there are unobserved individual characteristics common to both entry into VHC and subsequent health care spending.
  - Versions allow for controls on observed characteristics determining entry into VHC ('*propensity score matching*').
- *The Difference-in-Differences Estimator*
  - A non-parametric method.
  - Calculates the change in spending from before and after the decision month, and compares the average change for VHC individuals with that for non-VHC individuals.
  - Controls for unobserved individual characteristics that are constant over time (e.g., 'frailty').
  - No controls on observed characteristics.
- *The Difference-in-Differences Matching Estimators*
  - A non-parametric method.
  - Adds matching to the difference-in-differences estimator and so controls for observed characteristics determining entry into VHC.

1. Heckman Selection Estimator

A simple estimator of α is obtained by estimating (1) by ordinary least squares (OLS), using individuals both in and not in VHC, in some period following the month in which the people in VHC actually commenced in VHC. However, the estimator is biased if there is correlation between the VHC dummy variable d and the error U, that is, if entry into VHC is not random (see equation (6)).

The Heckman selection estimator assumes a form for the joint behaviour of d and U, via that for V and U; and essentially uses the form of the density to control for the correlation between d and U. This control takes the form of an additional term, the inverse Mills ratio, which is added to (1). OLS then gives the estimator of α.

From (1) and (3), and assuming joint normality of V and U, gives,

$$(11) \qquad \mathrm{E}[Y_{it}|d_i = 1] = \mathrm{E}[X_{it}\,\beta] + \alpha + \mathrm{E}[U_{it} \mid d_i = 1] = X_{it}\,\beta + \alpha + \rho\frac{\phi(Z_{it}\gamma)}{\Phi(Z_{it}\gamma)}$$

$$(12) \qquad \mathrm{E}[Y_{it}|d_i = 0] = \mathrm{E}[X_{it}\,\beta] + \mathrm{E}[U_{it} \mid d_i = 0] = X_{it}\,\beta - \rho\frac{\phi(\mathrm{Z}_{\mathrm{it}}\gamma)}{1-\Phi(\mathrm{Z}_{\mathrm{it}}\gamma)}$$

where ϕ is the probability density function of the standard normal, Φ is the cumulative distribution function of the standard normal, ρ depends on the correlation between V and U, $t > k$, and $\frac{\phi(\mathrm{Z}_{\mathrm{it}}\gamma)}{\Phi(\mathrm{Z}_{\mathrm{it}}\gamma)}$ is the inverse Mills ratio. OLS of Y_{it} on the appropriate right hand side in (10) and (11) gives the estimate of α.

An alternative estimator is one based on differences. From (1) and (2) and for $t > k$ and $t' < k$,

$$(13) \qquad Y_{it} - Y_{it'} = (X_{it} - X_{it'})\,\beta + d_i\,\alpha + \mu_{it} - \mu_{it'},$$

where the individual effect φ_i has been cancelled out. If μ_{it} and $\mu_{it'}$ are uncorrelated with V_{ik}, then no sample selection correction is needed. In the individual effects model, it is reasonable to expect that the correlation between $(\mu_{it} - \mu_{it'})$ and V_{ik} is less than that between U_{it} and V_{ik} , so that the bias from ignoring the sample selection may be smaller.

In this approach, complications such as the spending data being censored and there being outliers in the data are handled parametrically. For example, the long-tails can be handled by modelling the natural logarithms of the data and the censored nature of the data can be handled by replacing (1) and (2) by Tobit models. If Y_{it} is censored, then from Amemiya (1985, equation 10.4.6),

$$(14) \qquad \mathrm{E}[Y_{it}|\, Y_{it} > 0] = X_{it}\,\beta + \alpha\mathrm{E}[d_i \mid Y_{it} > 0] + \sigma\frac{\phi(X_{\mathrm{it}}\,\beta)}{\Phi(X_{\mathrm{it}}\,\beta)}$$

where σ is the standard deviation of U_{it}. In some of the spending categories there are relatively few observations with non-zero spending, e.g., private hospitals, which is also the category with the highest average spending.

A similar expression to (14), but without the term involving d_i, is obtained for $\mathrm{E}[Y_{it'}| Y_{it'}>0]$. Hence, in a model on $Y_{it} - Y_{it'}$ involving positive values for Y_{it} and $Y_{it'}$, the inverse Mills ratio terms almost cancel out and so can be ignored.

Note also that

$$\mathrm{E}[d_i \mid Y_{it} > 0] = \mathrm{P}(d_i = 1 \mid Y_{it} > 0) = \mathrm{P}(Y_{it} > 0 \mid d_i = 1)\mathrm{P}(d_i = 1)/\mathrm{P}(Y_{it} > 0),$$

which introduces further terms.

If the Mills terms are ignored then the difference $\mathrm{E}[Y_{it}| Y_{it} > 0] - \mathrm{E}[Y_{it'}| Y_{it'}>0]$ is equal to $\alpha\mathrm{E}[d_i \mid Y_{it} > 0]$. Since d_i takes on values zero and one, $\mathrm{E}[d_i \mid Y_{it} > 0] < 1$ and the impact of VHC is less than α. This comes about because, for some observations, adding α to the model still gives a censored observation – the observed change in spending is zero. However, if the quantity of interest is the average change in spending, then the smaller quantity, $\alpha\mathrm{E}[d_i \mid Y_{it} > 0]$ is the appropriate one.

Replacing the conditional expectation in (14) with an unconditional expectation gives

$$(15) \qquad \mathrm{E}[Y_{it}] = \Phi(X_{it}\,\beta)\left\{X_{it}\,\beta + \alpha\mathrm{E}[d_i \mid Y_{it} > 0] + \sigma\frac{\phi(X_{\mathrm{it}}\,\beta)}{\Phi(X_{\mathrm{it}}\,\beta)}\right\}$$

The model is no longer linear.

Separating the observations into those corresponding to individual in VHC and not in VHC leads to expressions like $\mathrm{E}[Y_{it}| Y_{it}>0, d_i = 1]$ or $\mathrm{E}[Y_{it}| d_i = 1]$, which again lead to Mills-like terms.

A version of the Heckman procedure was applied in our initial report (Access Economics, August 2002). The estimator depends on the assumed form of the joint behaviour of V and U. In that report, time constraints meant that we were unable to check the sensitivity of the results to our assumptions; we used the same model for all spending categories; etc. The analysis here shows some of the ways in which it was an approximation to a more complicated model.

Matching is based on expressions like (5), with no parameterisation of the model. But it is still appropriate to make explicit the assumptions it uses, and to see how it handles selection, censoring, etc.

2. The Matching Estimator

The aim in matching is to develop a proxy for the conditions of an experiment. Matching is more general than the Heckman estimator in that no particular specification has to be assumed. However, it rests on assumptions about the selection into VHC and has heavy data requirements.

For each individual in VHC, matching finds one or more non-VHC individuals with similar values of the X variables. The difference in average spending between the VHC individuals and non-VHC matched individuals is taken as the estimate of α. The estimator therefore relies on the assumption that the non-VHC individuals behave in the same way as the VHC individual would if he or she were not treated. In other words, the selection into VHC is independent of the spending, after the X's have been taken into account – there is no correlation between U and V. More formally, as in equation (10),

$$\begin{aligned}\alpha_T &= E[Y_{it}^T \mid d_i = 1, X] - E[Y_{it}^C \mid d_i = 1, X] \\ &= \{E[Y_{it}^T \mid d_i = 1, X] - E[Y_{it}^C \mid d_i = 0, X]\} - \{E[Y_{it}^C \mid d_i = 1, X] - E[Y_{it}^C \mid d_i = 0, X]\}.\end{aligned}$$

The first term in {} on the right-hand-side of the last expression is estimated in the matching. The second term is zero if, conditional on X, Y_{it}^C is independent of d_i. If there is correlation between Y_{it}^C and d_i, i.e., between entry into VHC and subsequent spending, then the matching estimator may be biased.

Essentially, matching plays the role of the X variables in (1). Instead of controlling for the X variables, via a regression, the method matches on the X's.

If matching is done on the X's, then the method is highly data and computer intensive. A way around this is to use *propensity score matching*. This is matching done on the basis of the propensity to enter into VHC, given the characteristics X (and Z): $P(X) = P(enter\ VHC|X) = P(d = 1|X)$. A logit or probit model can be used to estimate $P(X)$, essentially estimating equation (3). Rosenbaum and Rubin (1983 and 1984) give conditions under which matching on $P(X)$ can be used in place of matching on X. The matching is then one-dimensional and relatively straightforward.

We use probit models to estimate the $P(X)$.

Matching assumes that $0 < P(d=1|X) < 1$ in order to ensure that all VHC individuals have counterparts in the non-VHC population. For example, if the probability was equal to one, then all individuals with those X's would be in VHC and there would be no non-VHC individuals for available matching. If there are values of X for which no non-VHC individuals are available, then they are omitted from the analysis and the conclusion is qualified.

3. The Difference-in-Difference Estimator

The difference-in-difference (DID) estimator is based on the change in spending between the months before the decision month and after the decision month. The DID estimator compares the change in VHC individuals with the change in non-VHC individuals:

$$(16) \qquad \hat{\alpha} = (\overline{Y}_t^{(1)} - \overline{Y}_{t'}^{(1)}) - (\overline{Y}_t^{(0)} - \overline{Y}_{t'}^{(0)})$$

In a world with no changes through time, other than individuals entering VHC, the two averages in the second term on the right hand side bounded by parentheses should be similar and the second term will be approximately zero. The first term in parentheses on the right-hand-side may include changes due to VHC, so that $\hat{\alpha}$ estimates the effect of VHC. In the linear model, this difference targets the parameter α. In the non-linear models, it simply targets the corresponding difference in means of Y.

Alternatively, suppose that elements of change, e.g, policy changes, seasonality or price changes, affect all individuals. The change for non-VHC individuals is a value *x*, The change for VHC individuals includes both *x* and the impact of VHC, α. The difference-in-differences is $x + \alpha - x = \alpha$ – the effect of VHC.

Notice that because the estimator is based on differences, if the error terms are additive as in (7) and (8), *constant* elements in the unobservable *U* have no effect – the φ_i in (4) cancel out. For example, if VHC individuals are 'frailer' than non-VHC individuals and the 'frailty' leads to higher spending, then we expect higher average spending for VHC individuals, both before and after they enter VHC. But all that matters is how the spending changes. Suppose that VHC decreases 'frailty'. Then it will decrease spending. For the non-VHC individuals, spending does not change, and so the estimate of α is the average fall for VHC individuals.

Common seasonal and trend factors cancel out.

If the *non-constant* elements of U, the μ_{it} in (4), are correlated with d, then the estimator is biased. Furthermore, the estimator does not control for the X's, so individual specific X's correlated with d can also bias the estimator. For example, suppose that spending temporarily increases just before the decision month, and this increases the probability of entry into VHC. Subsequent decreases in spending are attributed to VHC.

The last example suggests that control on the X's will be helpful. This leads to the next estimator.

4. The Difference-in-Differences Matching Estimator

The estimator is based on difference-in-differences between VHC individuals and a set of non-VHC individuals obtained by propensity score matching. As in the DID estimator, it is robust to constant elements of U if the errors are additive. It can also handle a case such as the example in the previous section involving a temporary increase in spending. The VHC individual with the temporary increase would be matched with a similar non-VHC individual. Both will have a subsequent decrease in spending and the difference will be close to zero.

The estimator may still be biased if there are elements of the unobservable U that are correlated with the entry into VHC.

2.2 CHOICE OF ESTIMATION METHOD

Our assessment, based on the experience gained from the analysis done for the initial report and the additional analysis done for this report, is that the matched DID estimator is the preferred approach. The time available for this consultancy also means that we were able to implement this method. This is not an exaggeration since the computer runs for the basic model, over all classes of individuals and categories of spending, took over 50 hours.

2.3 IMPLEMENTATION OF MATCHING

Matching

The general form of the matching estimator is

$$(17) \qquad \hat{\alpha}_{MM} = \frac{1}{N_{VHC}} \sum_{i \in VHC} \left\{ Y_{it} - \sum_{l \notin VHC} W_{il} Y_{lt} \right\}$$

where W_{il} is a weight placed on the comparison observation l for individual i, N_{VHC} is the number of individuals entering VHC in decision month k, and $i \in (\notin) VHC$ means that individual i is in (not in) VHC.

In the matched DID procedure, Y_{it} is replaced by $(Y_{it} - Y_{it'})$, where Y_{it} is post-entry spending and $Y_{it'}$ is pre-entry spending:

$$(18) \qquad \hat{\alpha}_{MMDID} = \frac{1}{N_{VHC}} \sum_{i \in VHC} \left\{ (Y_{it} - Y_{it'}) - \sum_{l \notin VHC} W_{il} (Y_{lt} - Y_{lt'}) \right\}$$

We consider two forms of matching, corresponding to different set of weights, W_{il}.

Nearest Neighbour Matching

The comparison observation is the non-VHC observation closest to Y_{it} in terms of the value of the propensity score, $P(X)$. In (4), $W_{il} = 1$ for the nearest neighbour and $= 0$ otherwise.

Kernel Matching

The comparison group for each VHC individual is made up of all non-VHC individuals, although the weights applied to the non-VHC individuals decrease as the differences in propensity scores increases. The aim of the kernel marching is to decrease the variance of the estimator. Kernel matching has been used in many recent studies.

We implement kernel matching using the Epanechnikov kernel with bandwidth given by,

$$(19) \quad h = 0.5\, A\, n^{-1/5}$$

where n is the number of non-VHC individuals in the comparison group and A is the standard deviation of the propensity scores of the non-VHC individuals.

It was noted above that for VHC individuals with extreme X and hence $P(X)$ values, there may be no non-VHC individual with a similar $P(X)$ value. Individuals with no matches are excluded from the analysis and the conclusions only apply to the observations for which matches can be found. This is similar to dropping X outliers.

Standard Errors

The complicated and non-parametric nature of the DID and matching estimators implies that standard errors are difficult to obtain by analytical methods. Many studies employ bootstrap methods to obtain standard errors. But bootstrap methods are extremely computer intensive in large data sets and we did not have time to obtain such estimates. We base standard errors on simple approximations, such as the standard error of a mean being the standard error of the data divided by the square root of the sample size. This is an approximation because it ignores the fact that the random variables used to form the mean are themselves based on estimated quantities, e.g., the estimation of the propensity score model.

2.4 Model selection, testing, and sensitivity

In this section, we discuss the procedures used to specify the propensity score model, the tests applied to the model, and the procedures applied to check the sensitivity of the model to the assumptions.

2.4.1 Choice of propensity score model

The explanatory variables in the propensity score models include the age, the previous two months spending in the seven categories, the State/Territory of residence in the decision month, and the disability pension rate. Gender is taken into account by estimating separate models for males and females. Separate models are also estimated for ex-HACCs and non-transitionals. Only gold card holders are included in the analysis.

The assumptions behind the use of the propensity score in matching imply that the distribution of the *X*'s should be approximately the same for VHC and non-VHC individuals, after the propensity score has been taken into account. Dehejia and Wahba (2002) suggest using the following algorithm for testing this assumption:

1. Start with a parsimonious specification of the model.
2. Stratify all observations by the estimated propensity score, e.g. into bins based on score ranges (0-0.1,0.1-0.2,…,0.9-1).
3. Statistical Test: Test the difference in means of the *X* variables.
    a. If the *X*'s are balanced, in that there are few statistically significant differences between the means for VHC and non-VHC individuals, stop.
    b. If *X*'s are not balanced in some stratum, divide the stratum unto finer strata and reevaluate.
    c. If the *X*'s are not balanced for many strata, modify the model by adding extra terms.

We checked the specifications of the models using this algorithm.

In specifying the propensity model, we also consider the *t*-statistics on the variables, and the Akaike, Schwartz, and Hannan-Quinn model selection criteria,

Akaike information criterion (AIC)	$-2(l/T) + 2m/T$
Schwartz criterion (SC)	$-2(l/T) + m\log(T)/T$
Hannan-Quinn criterion (HQ)	$-2(l/T) + 2m\log(\log(T))/T$

where l is the log-likelihood, m is the number of X variables in the model (including the intercept) and T is the number of observations (e.g., Judge et al. 1985, section 7.5.2).

2.4.2 Changes in propensity score models and X's over time

Of interest is whether the characteristics of the individuals selected into VHC have changed over the period January 2001 – May 2002. For example, were the early entrants older or sicker, as measured by average spending, than the late entrants? Were they concentrated in different States/Territories? There are two components to this – did the characteristics change, and did the propensity score models change?

We tracked the means of the X variables over time and compared the parameters in the propensity score models in different decision months.

2.4.3 Choice of comparison group

For individuals who enter into VHC, the month of entry is known. Each of the estimation methods requires a set of non-VHC individuals who did not enter into VHC in that month – the 'zeros' in the propensity score model and the comparison group in the Heckman and matching estimators. It is not known which individuals were considered for VHC by the assessors or by their health care providers or others, so the non-entry group must be picked by some other method.

As noted in our earlier report, using all non-VHC individuals as the comparison group implies small probabilities/propensities for entering VHC. For example, if 8,000 out of 300,000 entered in a particular month, then the average probability is approximately 0.027. This makes it difficult to identify any sample selection effects in the Heckman procedure. In the estimated models, it also implies a relatively small difference in the average probability of entry for VHC and non-VHC individuals (e.g., 0.08 versus 0.025 in one model). In other words, there is relatively little to distinguish the VHC and non-VHC individuals. Of course, this may simply reflect reality; but it is also unlikely that all non-VHC individuals are considered for entry in a particular month. For example, many have no contact with the 'system' in any particular month. Using all the individuals also implies increased computation times.

However, we do not try to model the factors that might lead an individual to be considered for VHC. Instead, we use a randomly selected comparison group.

For each decision month, the total comparison group (males and females) contains around 7,000 individuals – the individuals are randomly chosen with probability based on groups of size 7,000. An individual may appear in more than one comparison group, but individuals already in VHC or who entered VHC in the five months following the decision month are excluded. The check of the last condition is applied after the 7,000 have been selected, so that the actual comparison groups are smaller than 7,000. The total comparison group is divided into males and females for the respective models.

Relative to a choice of, say, excluding all individuals who entered VHC up to and including May 2002, the choice of five months for the exclusion period implies that, for the early decision months, the estimates for the out-months are partly based on some VHC individuals. Excluding these individuals would leave out any 'spikes' in their spending in the months before they enter VHC. This would reduce the spending of the comparison group and increase the estimate of the impact of VHC. Conversely, the five month choice implies that estimates for in-months for the late decision months are not affected by individuals who entered VHC in June and July 2002.

2.4.4 Implications of choice of comparison group

It is not meaningful to consider the implications of the choice of a randomly selected comparison group relative to having the 'true' comparison group, since such a group may not even exist. Better comparisons are with alternative randomly selected groups and with all the available zeros in the dataset.

With respect to a different randomly selected comparison group, the issue is the variance of the estimate of the impact across different groups, assuming that the estimates from all groups are estimating the same quantity, the impact of VHC relative to a randomly selected comparison group. This is one area in which the standard errors on the estimates do not reflect all the sources of randomness. A different comparison group would imply different estimates of the propensity models (a different draw of the estimates from their sampling distribution), and the estimation of the propensity model is not taken into account in the standard errors.

Next, consider the comparison with all the available zeros. First note that in the matching, with a given set of propensity scores, using the random sample rather than all the zeros does not imply a bias. The issue is again one of variance. The effects of using the random sample come through the estimation of the propensity model.

In the estimation of the propensity model, the randomly selected group can be thought of as a choice-based sample (e.g., Amemiya, 1985, section 9.5). Relative to the non-VHC individuals, the VHC individuals are over-sampled. Estimating the model using a likelihood function reflecting the choice-based sample would make the procedure comparable to estimating the model on the entire set of zeros. Estimating the propensity model using all the zeros, and including all the zeros in the matched DID estimator, would be estimating the impact of VHC relative to all non-VHC individuals. This may be different from the impact relative to a randomly selected group.

Not taking the choice-based sampling into account in the propensity model implies the estimators of the parameters in the propensity model are biased relative to those in the model will all zeros. The propensity scores may be different and the matched DID estimator may be estimating the different impact.

Our choice of using a randomly selected comparison group implies that we are estimating the impact of VHC relative to a randomly selected comparison group, rather than relative to all non-VHC individuals.

Finally, in our proposal we referenced the work of Smith and Todd (2000), who note that the use of the odds ratio, $P/(1-P)$, in the matching corrects for the choice-based sampling. But again, this would be equivalent to using all non-VHC individuals, which we are not doing. Hence, we weight using P.

2.5 DATA ISSUES

a) One quarter sized sample

As in our earlier analysis, we estimate the results using a one quarter sized sample of the available data. More specifically, we include all VHC individuals and let non-VHC individuals enter into the sample with probability ¼. The aim of the sampling is to decrease computation times.

b) Treatment of the Decision Month

Data from the decision month is excluded from the analysis, since it is unclear whether to attribute the spending to pre-entry or post-entry. Furthermore, the decision month often contain short-term spending effects. Including these in spending comparisons could give misleading results. A similar argument applies to pre-decision month spending, as detailed in the next point.

c) Treatment of Pre-Decision Month Spending

As noted above, it turns out that spending from the two months immediately prior to the decision month, but not before that, are included in the propensity score model (and hence in the matching). Another way to say this is that given that these two months are in the model, spending from earlier months are not significant in explaining entry into VHC.

However, data from these two months are not included in the estimates of the impact of VHC. As is well known, average spending for entrants tend to 'spike' just before entry; and the matching implies a similar pattern for the matched non-VHC individuals. However, the spending does not line up exactly (recall also the discussion of private hospital spending). Hence, we exclude those months.

The comparison point for post-decision month spending is the average spending in months 3 through 8 before the decision month.

d) Treatment of Young Individuals

As in our earlier analysis, we exclude individuals who were aged 50 years or less in January 2001.

e) Postcode Data

The postcode of each individual, at the time of spending, is available in the dataset. These were mapped into DVA regions via postcode-SLA and SLA-region concordances. Some observations were lost due to recent postcode changes and mismatches in the two concordances – the SLA-region concordance pre-dates some recent changes to SLA definitions. Postcodes were also mapped into codes for capital cities, other metropolitan centres, large rural centres, small rural centres, other rural areas, remote centres, and other remote areas.

When a postcode area was split between more than one SLA, the individual was randomly assigned to the SLA according to the split of the population in the postcode area between the SLAs.

f) Missing Data

In general observations with missing data were deleted.

g) Late Entrants into the Population

For individuals who received a gold card late in the sample period, earlier spending is recorded as missing.

h) Deaths

Individuals who died during the treatment period are excluded from the matched DID estimator, although they are included in the propensity score model. The exclusion (rather than including them up until the month of death) is to speed the computations.

References

a) Referenced in report:

Access Economics (August 2002), Veterans' Home Care Programme: Analysis of Impact on Other Programmes of the Department of Veteran' Affair. In-confidence report.

Amemiya, T, (1985), *Advanced Econometrics*, Harvard University Press, Cambridge.

Dehejia, R.H. and S. Wahba (2002), 'Propensity Score-Matching Methods for Nonexperimental Causal Studies,' *Review of Economics and Statistics*, 84, 151-161.

Judge. G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T-C Lee (1985), *The Theory and Practice of Econometrics*, second edition, John Wiley and Sons, New York..

Rosenbaum, P. and D. Rubin (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects,' *Biometrika*, 70, 41-55.

Rosenbaum, P. and D. Rubin (1985), 'Constructing a Control Group using Multivariate Matched Sampling Methods that Incorporate the Propensity Score,' *American Statistician*, 39, 33-38.

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Smith, J. and P. Todd (2000), 'Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators,' working paper, University of Western Ontario.

Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

b) Additional background references used in preparing this report:

Blundell, R. and M. Costa Dias (2000), 'Evaluation Methods for Non-Experimental Data,' *Fiscal Studies*, 21, 427-468.

Heckman, J.J., H. Ichimura and P.E. Todd (1997), 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,' *Review of Economic Studies*, 64, 605-654.

Heckman, J.J., H. Ichimura and P.E. Todd (1998), 'Matching as an Econometric Evaluation Estimator,' *Review of Economic Studies*, 65, 261-294.

Heckman, J.J., H. Ichimura, J. Smith, and P. Todd (1998), 'Characterizing Selection Bias using Experimental Data,' *Econometrica*, 66, 1017-1098.

Heckman, J.J. and P. Todd (1995), 'Adapting Propensity Score Matching and Selection Models to Choice-based Samples,' manuscript, University of Chicago.

Maddala, G.S. (1983), *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.