

Unidimensional Latent Regression

Ray Adams and Margaret Wu, 29 August 2010

The term latent regression refers to the direct estimation of regression models from item response data. To illustrate the use of latent regression, consider the following typical situation. There are two groups of students, group A and group B, and it is of interest to estimate the difference in the mean achievement of the two groups. A common approach is to administer a test to the students and then use this test to produce achievement scores for all of the students. A standard procedure can then be applied, such as regression (which, in this simple case, becomes identical to a *t*-test), to examine the difference in the means of the achievement scores. Depending upon the model that is used to produce 'student scores,' this approach can result in misleading inferences about the differences in the means. Using the latent regression methods described by Adams, Wilson and Wu (1997), ConQuest avoids such problems by directly estimating the difference in the mean achievement of the groups from the item response data without first producing individual student scores.

A LATENT VARIABLE T-TEST

The data used here are a subset of the data that were collected by Lokan, Ford and Greenwood (1996) as part of the Third International Mathematics and Science Study (TIMSS) (Beaton *et al.* (1996)). The TIMSS data that we will be using are the mathematics achievement test data, collected from a sample of 6800 students in their first two years of secondary schooling in Australia¹.

The TIMSS study used a sophisticated test item rotation plan that enabled achievement data to be gathered on a total of 158 test items while restricting the testing time for any individual student to 90 minutes. Details on how this was achieved are described in Adams and Gonzalez (1996). In this tutorial, we will be using the data to examine grade differences and gender differences in students' mathematics achievement as tested by the TIMSS tests.

The data set used in this sample analysis, `ex5.dat`, contains 6800 lines of data, one line for each student that was tested. Columns 20 to 176 contain the item responses. The TIMSS tests consist of multiple choice, short answer and extended response questions. For the multiple choice items, the codes 1, 2, 3, 4 and 5 are used to indicate the response alternatives to the items. For the short answer and extended response items, the codes 0, 1, 2 and 3 are used to indicate the student's score on the item. If an item was not presented to a student, the code . (a period) is used; if the student failed to attempt an item and that item is part of a block of non-attempts at the end of a test, then the code R is used. For all other non-attempts, the code M is used. The first 19 columns of the data set contain identification and demographic information. In this example, only the data in columns 17 through 19 are used. Column 17 contains the code 0 for male students and 1 for female students; column 18 contains the code 0 for lower grade (first year of secondary school) students and 1 for upper grade (second year of secondary school) students; and column 19 contains the product of columns 17 and 18, that is, it contains 1 for upper grade female students and 0 otherwise.

¹ These 6800 students were randomly selected from a larger Australian TIMSS sample of over 13 000 students in their first two years of secondary schooling.

Unidimensional Latent Regression

In the first sample analysis that uses these data, it is of interest to estimate the difference in achievement between the lower and upper grades. To illustrate the value of directly estimating the differences using latent regression, only the first six items are used. Later in the tutorial, we will compare the results obtained from analysing only these six items with the results obtained from analysing all 158 items.

The files used in this first sample analysis are:

ex5a.cqc	The command statements for the first analysis.
ex5.dat	The data.
ex5.nam	The variable labels for the items.
ex5a.mle	Maximum likelihood ability estimates for the students.
ex5a.eap	Expected a-posterior ability estimates for the students.
ex5a.shw	Selected results of the analysis.
ex5a.itn	The results of the traditional item analyses.

(The last four files will be created when the command file is executed.)

The command file used in this sample analysis is shown in Figure 1.

```
1.  datafile ex5.dat;
2.  title    Australian TIMSS Mathematics Data--First Six Items;
3.  format  gender 17 level 18 gbyl 19
      responses 20-25;
4.  labels  << ex5.nam;
5.  key    134423 ! 1;
6.  regression level;
7.  model  item;
8.  estimate !fit=no;
9.  show cases ! estimates=mle >> ex5a.mle;
10. show cases ! estimates=eap >> ex5a.eap;
11. show ! tables=3 >> ex5a.shw;
12. itanal >> ex5a.itn;
```

Figure 1 Sample Command File for a Latent Variable *t*-Test (Six Items)

1. Indicates the name and location of the data file.
2. Gives a title for this analysis. The text that is given after the command `title` will appear on the top of any printed output. If a title is not provided, the default, `ConQuest: Generalised Item Response Modelling Software`, will be used.
3. The `format` statement describes the layout of the data in the file `ex5.dat`. This format indicates that a code for gender is located in column 17, a code for level is located in column 18, column 19 contains the code for a variable we have called `gbyl`, and responses are to be read from columns 20 through 25. We have not given a name to the responses, so they will be referred to as `item`.
4. A set of labels for the items are to be read from the file `ex5.nam`.

NOTE: *The file `ex5.nam` contains labels for all 158 items. These are all read and held in memory by ConQuest, even though we are only using the first six items in this analysis.*

5. The argument of the `key` statement identifies the correct response for each of the six multiple choice test items. In this case, the correct answer for item 1 is 1, the correct answer for item 2 is 3, the correct answer for item 3 is 4, and so on. The length of the key statement argument is six characters, which is the length of the response block given in the `format` statement. The `key` statement option indicates that each correct answer will be recoded to 1. By default, incorrect answers will be recoded to 0.

NOTE: *These data contain three kinds of missing-response data. The codes for these missing-response data are `.` (a period), `M`, and `R`. In this analysis, ConQuest will treat `.` as missing-response data, since it is one of the default missing-response codes. Those data coded `M` and `R` will be treated as incorrect, because these codes do not match the values in the `key` statement argument.*

6. The independent variables that we want to include as predictors of the latent variable are included as arguments in the `regression` statement. By including the variable `level` as the argument here, we are instructing ConQuest to regress latent ability onto `level`; and in this case, since `level` is coded 0 (lower grade) and 1 (upper grade), ConQuest will estimate the difference between the means of these two groups. The `regression` statement is used to describe the ConQuest population model.

TIP: *If you want to regress the latent variable onto a categorical variable, then the categorical variable must first be appropriately recoded. For example, dummy coding or contrast coding can be used. A variable used in regression must be a numerical value, not merely a label. For example, gender would normally be coded as 0 and 1 so that the estimated regression is the estimated difference between the group means. Remember that the specific interpretation of the latent regression parameters depends upon the coding scheme that you have chosen for the categorical variable.*

7. The `model` statement here contains only the term `item` because we are dealing with single-faceted dichotomous data.
8. The `estimate` statement is used to initiate the estimation of the model. The `fit=no` option is included because in this sample analysis we are not concerned with the item fit and it will save time if the fit statistics are not computed.
9. The `show` statement produces a display of the results from fitting the model. Here the `cases` argument is used to request a set of ability estimates for the students. The `estimates=mle` option indicates that maximum likelihood estimates of the ability are requested, and they are redirected to the file `ex5a.mle`. When case estimates are requested, both the option indicating the type of estimate and redirection to a file are required.

Unidimensional Latent Regression

10. As for line 9, only we are requesting expected a-posteriori ability estimates rather than maximum likelihood ability estimates be written to the file `ex5a.eap`. In 'Latent Estimation and Prediction' in Chapter 12 of Wu, Adams, Wilson and Haldane (2007), the difference between these two types of ability estimates is described.
11. This third `show` statement writes the third results table to the file `ex5a.shw`. This table contains the parameter estimates for the population model.
12. The `itanal` statement produces some traditional item statistics and writes them to the file `ex5a.itn`.

RUNNING THE T-TEST SAMPLE ANALYSIS

To run this sample analysis, start the gui version of ConQuest and open the control file

```
Ex5a.cqc
```

Select Run -> Run All. ConQuest will begin executing the statements that are in the file `ex5a.cqc`; and as they are executed, they will be echoed in the Output window. When ConQuest reaches the `estimate` statement, it will begin fitting Rasch's simple logistic model to the data; as it does so, it will report on the progress of the estimation. This particular sample analysis will take 34 iterations to converge. Figure 2 shows an extract of the information that is reported as ConQuest iterates to a solution. The first regression coefficient is the constant, and the second is the regression coefficient of the variable `level` in the regression of latent ability onto `level`.

Figure 3 shows the contents of the file `ex5a.shw`. The values reported here are the parameter estimates for the population component of the ConQuest model—in this case, a regression of the latent ability onto grade level. In these data, the `level` variable was coded as 0 for the lower grade and 1 for the upper grade, so the results shown in Figure 3 indicate that the estimated mean of the lower grade is 0.671 and the mean of the upper grade is 0.231 higher (mean of higher grade=0.902). The conditional variance in the latent variable is estimated to be 1.207. If an item response model is fitted without the regression variable, the estimated mean and variance of the latent ability are 0.80 and 1.219 respectively.²

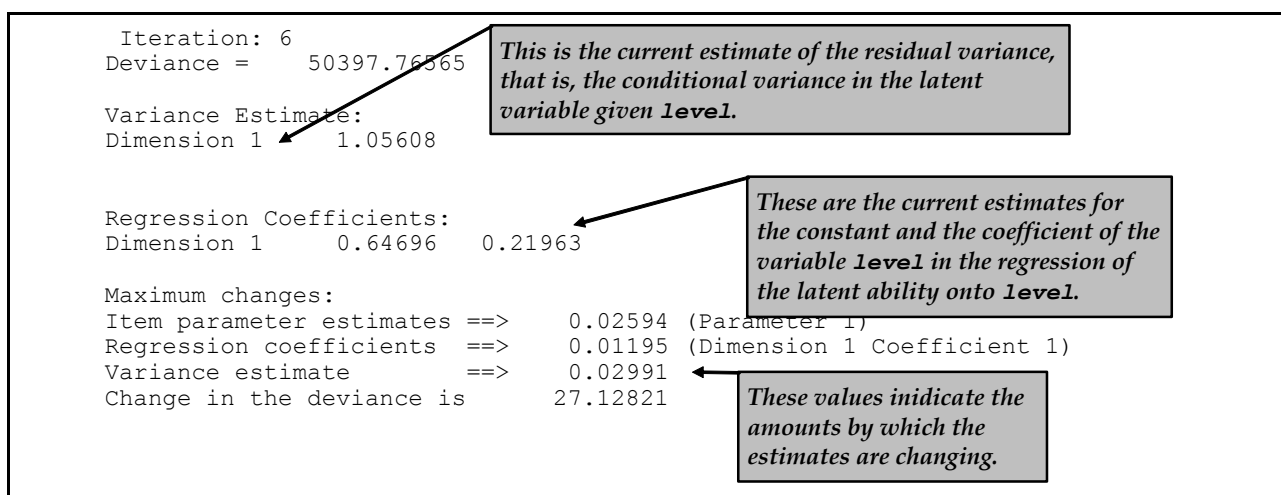


Figure 2 Reported Information on Estimation Progress for `ex5a.cqc`

² The current version of ConQuest does not report standardised regression coefficients or standard errors for the regression parameter estimates. Plausible values can be generated (via `show cases !estimates=plausible`) and analysed to obtain estimates of standard errors and to obtain standardised regression coefficients.

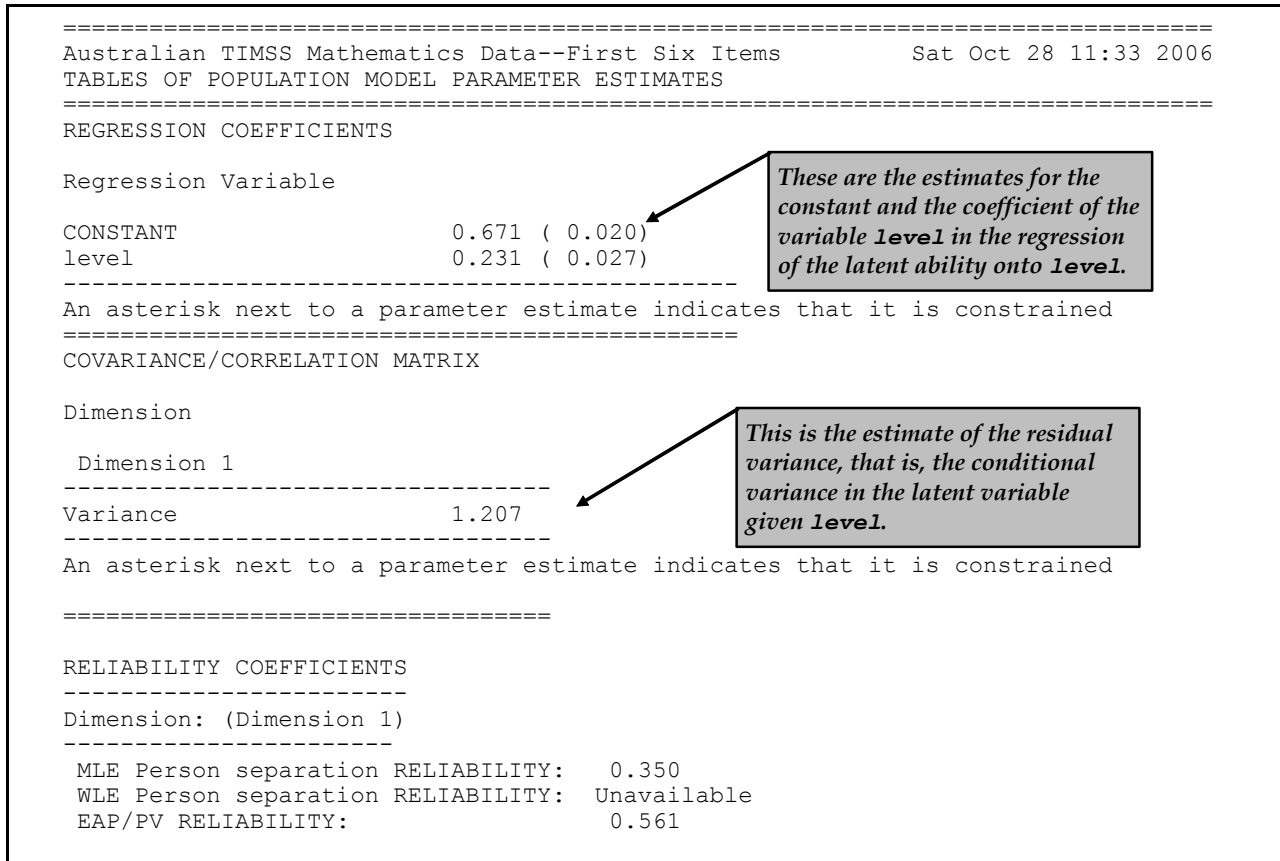


Figure 3 Population Model Parameter Estimates

The command file listed in Figure 1 also produces the files `ex5a.mle` and `ex5a.eap`. These files contain latent ability estimates for each of the 6800 students in the file `ex5.dat`. The format of these files is as follows. The file `ex5a.mle` contains one line of data for each student in the sample who provided a valid response to at least one of the six items that we have analysed – in this sample, 6778 students. Columns 1 through 5 contain an identification number for the case, which is the sequence number of the student in the original data file. Columns 6 through 15 contain the total score that the student attained, columns 16 through 26 contain the maximum possible score that the student could have attained, columns 27 through 37 contain the maximum likelihood estimate of the student’s latent ability, and columns 38 through 48 provide an asymptotic standard error for that ability estimate. An extract from `ex5a.mle` is shown in Figure 4.

	1	2	3	4	
	123456789012345678901234567890123456789012345678				(col. nos.)
1	3.00	6.00	-0.00648	0.84976	
2	4.00	6.00	0.74404	0.89598	
3	1.00	6.00	-1.71624	1.10927	
4	0.30	6.00	-3.08143	1.88317	
5	3.00	6.00	-0.00648	0.84976	
6	5.00	6.00	1.72089	1.11607	
7	5.00	6.00	1.72089	1.11607	
8	1.00	6.00	-1.71624	1.10927	
9	3.00	6.00	-0.00648	0.84976	
.	
.	

Figure 4 An Extract from the MLE File `ex5a.mle`

EXTENSION: *The maximum likelihood estimation method does not provide finite latent ability estimates for students who receive a score of zero or students who achieve the maximum possible score on each item. ConQuest produces finite estimates for zero and maximum scorers by estimating the abilities that correspond to the scores r and $M-r$ where M is the maximum possible score and r is an arbitrarily specified real number. In ConQuest, the default value for r is 0.3. This value can be changed with the `set` command argument `zero/perfect= r` .*

The file `ex5a.eap` contains three lines of data for each student in the sample who provided a valid response to at least one of the six items that we have analysed – in this case, 20 334 lines. The first line contains an identification number, which is the sequence number of the student in the original data file. The second line contains the expected value of the student’s posterior latent ability distribution—the so-called EAP ability estimate. The third line is the variance of the student’s posterior latent ability distribution; this can be used as the error variance for the EAP ability estimate. An extract from `ex5a.eap` is shown in Figure 5.

1	0.25916	0.68392	0.61255
2	0.74681	0.71617	0.57515
3	-0.69164	0.72285	0.56718
4	-1.27239	0.80593	0.46197
5	0.25916	0.68392	0.61255
6	1.29044	0.75783	0.52427
7	1.29044	0.75783	0.52427
8	-0.69164	0.72285	0.56718
9	0.25916	0.68392	0.61255
10	-0.20316	0.68233	0.61434
11	-0.20316	0.68233	0.61434
12	0.25916	0.68392	0.61255
.			
.			

Figure 5 An Extract from the EAP File `ex5a.eap`

WARNING: *The maximum likelihood estimate is a function of the item response data only; as such, it is not influenced by the population model. The EAP estimates are a function of both the population model and the item response model, so a change in the population model will result in a change in the EAP estimates.*

COMPARING LATENT REGRESSION WITH OLS REGRESSION

If the file `ex5a.mle` is merged with the `level` variable for each case, it is possible to regress the maximum likelihood ability estimates onto `level`. Similarly, if the file `ex5a.eap` is merged with the `level` variable, a regression of EAP estimates onto `level` can be carried out. The results obtained from these two regression analyses can be compared (see Figure 3). For the purposes of this comparison, we have also fitted a model without any regressors and added the EAP ability estimates from this run to the file `ex5a.out`, which we have provided.³

The results of ordinary least squares (OLS) regressions of the various estimates of latent ability onto `level` are shown in Table 1.

Table 1 OLS Regression Results Using Alternative Latent Ability Estimates

<i>Analysis Method</i>	<i>Latent Ability</i>		<i>Regression Coefficients</i>		<i>Conditional Variance</i>	<i>R²</i>
	<i>Mean</i>	<i>Variance</i>	<i>Constant</i>	<i>Level</i>		
Estimator*						
MLE	0.83	2.22	0.70	0.236	2.201	0.006
EAP with Regressor	0.80	0.68	0.67	0.231	0.664	0.019
EAP without Regressor	0.80	0.67	0.73	0.128	0.669	0.006
Direct ConQuest Estimation	0.80	1.22	0.67	0.231	1.207	0.010

* These are the ability estimates computed by ConQuest and then used as dependent variables in the analyses.

The last row of the table contains the results produced directly by ConQuest. Theoretical and simulation studies by Mislevy (1984, 1985) and Adams, Wilson and Wu (1997) indicate that the ConQuest results are the 'correct' results. The results in the table show that the mean of the latent ability is reasonably well estimated from all three estimators. The slight overestimation that occurs when using the MLE estimator is likely due to the ad-hoc approach that must be applied to give finite ability estimates to those students with either zero or perfect scores. The variance is overestimated by the MLE estimator and underestimated by the two EAP estimators. The overestimation of variance from the MLE ability estimator results from the independent measurement error component (see Wright and Stone, 1979) and a slight 'outwards' bias in the MLE estimates (see Lord, 1983, 1984). The underestimation of variance

³ The file `ex5a.out` contains the EAP and maximum likelihood ability estimates merged with the `level` variable for the 6800 students. The file contains one line per student, and the fields in the file are sequence number, `level`, maximum likelihood ability estimate (fourth field in Figure 4), EAP ability estimate when `level` is used as a regression variable (third field in Figure 5) and EAP ability estimate when no regressor is used.

from the EAP ability estimators results from the fact that the EAP estimates are ‘shrunk’ (see Lord, 1983, 1984).

For the regression model, we note that MLE estimates are reasonably close to the ConQuest results, the EAP estimates produced with the use of the regressor give results the same as those produced by ConQuest, and the EAP estimates produced without the regressor overestimate the constant term and underestimate the level effect. As was the case with the means, the difference between the MLE-based estimates and the ConQuest-based estimates for the constant term is likely due to the ad-hoc treatment of zero and perfect scores when ConQuest generates the maximum likelihood point estimates. The EAP estimates produced with the use of the regressor give unbiased estimates of the regression coefficients, while the estimates produced with the EAP without regressor are shrunk. The conditional variances behave in the same fashion as the (unconditional) variance of the latent ability.

EXTENSION: *Plausible values enable the unbiased estimation of the parameters of any submodel of the population model that is specified in the ConQuest analysis and is used to generate the plausible values.*

None of the point estimators of students’ latent abilities can be relied upon to produce unbiased results for all of the parameters that may be of interest. This is particularly true for short tests, as is the case here. When tests of 15 or more items are used, both MLE and EAP estimators will produce results similar to those produced directly by ConQuest.

AVOIDING THE PROBLEM OF MEASUREMENT ERROR

The differences between the regression results that are obtained from ConQuest and from the use of ordinary least squares using the various point estimates of latent ability can be avoided by using longer tests. In Figure 6, we present the command file for a sample analysis that will read and analyse all of the items in the file `ex5.dat`.

The files that we use in this second sample analysis are:

<code>ex5b.cqc</code>	The command statements used for the second analysis.
<code>ex5.dat</code>	The data.
<code>ex5.nam</code>	The variable labels for the items.
<code>ex5b.prm</code>	Initial values for the item parameter estimates.
<code>ex5b.reg</code>	Initial values for the regression parameter estimates.
<code>ex5b.cov</code>	Initial values for the variance parameter estimates.
<code>ex5b.shw</code>	Selected results of the analysis.
<code>ex5b.eap</code>	Expected a-posterior ability estimates for the students.

(The last two files will be created when the command file is executed.)


```

1. datafile ex5.dat;
2. title Australian TIMSS Mathematics Data--All Items;
3. format gender 17 level 18 gbyl 19
   responses 20-176;
4. labels << ex5.nam;
5. codes 0,1,2,3,4,5,M;
6. recode (R) (.);
7. key
13442341153114122133113244235223341323324535
24322331211313511511324113143242223435112411
14141222133113123244122143223342135422133251
12253131111111111111111111111141!1;
8. key
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXX2X2XXXXX22XXXXXX2XX2XXXXXXXXXXXXXXXX2XXXX
XX2X2XXX2XXXXX2XXXXX2XX2XXXXXXXXXX2XXXXXXXXXXXX
XXXXXXXX222222222222222222X2!2;
9. key
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXX3X3XXXXX33XXXXXX3XX3XXXXXXXXXXXXXXXX3XXXX
XX3X3XXX3XXXXX3XXXXX3XX3XXXXXXXXXX3XXXXXXXXXXXX
XXXXXXXX333333333333333333X3!3;
10. regression level;
11. model item+item*step;
12. import init_parameters << ex5b.prm;
13. import init_reg_coefficients << ex5b.reg;
14. import init_covariance << ex5b.cov;
15. estimate !fit=no;
16. show !tables=3 >> ex5b.shw;
17. set p_nodes=1000;
18. show cases !estimates=eap >> ex5b.eap;

```

Figure 6 Sample Command File for a Latent Variable *t*-Test (157 Items)

- 1.-4. As for Figure 1, except the `title` statement has been changed to indicate all items are being analysed rather than the first six and the response block in the `format` statement has been enlarged to include all 157 responses.
5. In this analysis, we would like to treat the data coded R as missing-response data and the data coded M as incorrect. It is necessary therefore to make an explicit list of codes that excludes the R. This is in contrast to the previous sample analysis in which we did not provide a code list. In that case, all data in the file were regarded as legitimate, and those responses not matching a key were scored as incorrect.
6. Here the R code is recoded to . (period), one of the default missing-response codes. Strictly speaking, this recode statement is unnecessary since the absence of the R in the code string will ensure that it is treated as missing-response data. It is added here as a reminder that R is being treated as missing-response data.
- 7.-9. The `key` statement argument is now 157 characters long because there are 157 items. This test contains a mixture of multiple choice, short answer and extended response items, so we are using three `key` statements to deal with the fact that the short answer and extended response items are already scored. The first `key` argument contains the keys for the multiple choice items; and for short answer and extended response items, the code 1 has been entered. Any matches to this `key` argument will be recoded to 1, as

shown by the option. In other words, correct answers to multiple choice items will be recoded to 1; and for the short answer and extended response items, 1 will remain as 1. All other codes will be recoded to 0 (incorrect) after the last key statement and any recode statements have been read. The second and third key statements contain the character X for the multiple choice items and 2 and 3 respectively for the short answer and extended response items. As X does not occur in the response block of the data file, these key statements will have no effect on the multiple choice items (correct answers to which have been recoded to 1 by the first key statement), but the short answer and extended response items will have code 2 scored as 2 and code 3 scored as 3. While the second and third key statements don't change the codes, they prevent the 2 and 3 codes in the short answer and extended response items from being recoded to 0, as would have occurred if only one key statement were used.

10. As for line 6 of Figure 1.
11. This model statement yields the partial credit model. In the previous sample analysis, all of the items were dichotomous, so a model statement without the `item*step` term was used. Here we are specifying the partial credit model because it will deal with the mixture of dichotomous and polytomous items in this analysis.
- 12.-14. This analysis takes a considerable amount of time, so initial value files are used to import a set of starting values for the item, regression and variance parameter estimates.
15. In this sample analysis, we are not concerned with the properties of the items, so we are specifying the `fit=no` option to speed up the analysis.
16. The `set` command is used to alter some of ConQuest's default values. The `p_nodes=1000` argument requests that 1000 nodes be used when EAP estimates are produced and when plausible values are drawn. The default value for `p_nodes` is 2000. Reducing this to 1000 decreases the time necessary to compute EAP estimates.

EXTENSION: *ConQuest uses the Monte Carlo method to estimate the mean and variance of the marginal posterior distributions for each case. The system value `p_nodes` governs the number of random draws in the Monte Carlo approximations of the integrals that must be computed. In the unidimensional case, this method is inefficient; and future versions of ConQuest will implement a quadrature method as a more efficient alternative.*

WARNING: *For cases with extreme latent ability estimates, the variance of the marginal posterior distribution may not be estimated accurately if `p_nodes` is small. Increasing `p_nodes` will improve the variance estimates. On the other hand, for EAP estimates, moderate values of `p_nodes` are sufficient.*

17. This `show` statement writes a file containing the MLE ability estimate for each case.
18. This `show` statement writes a file containing the EAP ability estimate for each case.
19. This `show` statement writes the population model parameter estimates (table 3) to `ex5b.shw`.

RUNNING THE SECOND T-TEST SAMPLE ANALYSIS

To run this sample analysis, start the gui version of ConQuest and open the control file

Ex5b.cqc

Select Run -> Run All. ConQuest will begin executing the statements that are in the file ex5b.cqc; and as they are executed, they will be echoed in the Output Window. When ConQuest reaches the `estimate` statement, it will begin fitting the partial credit model to the data. In this case, only two iterations will be necessary because the initial values that were provided in the files `ex5b.prm`, `ex5b.reg` and `ex5b.cov` are the output of a full analysis that have been performed on a previous occasion.

Figure 7 shows the contents of `ex5b.shw`. A comparison of the results reported here with those reported in Figure 3 is quite interesting. Recall that the results in Figure 3 are from fitting a similar latent regression model to the first six items only – the set of items taken by all students. What we note is that the variance estimates are very similar, as is the regression coefficient for `level`. In fact, this similarity is quite remarkable given that the first analysis used only six of the 158 items and approximately one-fifth of the data that were actually available. The constant terms are quite different. The difference between the estimates for the constant is due to the model identification constraint. In the previous analysis, the item response model was identified by setting the mean difficulty of the first six items to zero. In this second run, the mean difficulty of all 158 items is set to zero.

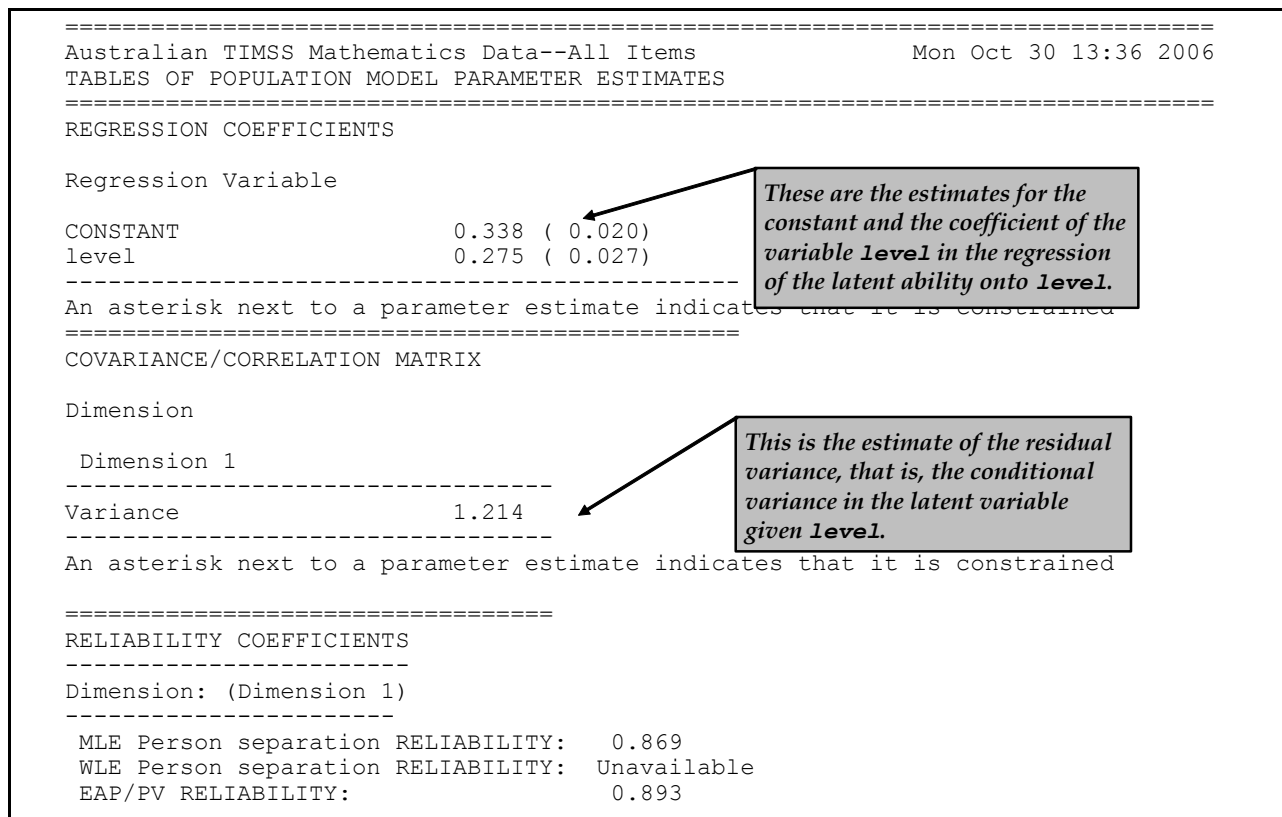


Figure 7 Population Model Parameter Estimates

COMPARING LATENT REGRESSION WITH OLS REGRESSION FOR THE SECOND SAMPLE ANALYSIS

As with the previous sample analysis, we produced a file of EAP ability estimates and then merged these with the `level` variable for each case. For the purposes of this comparison, we have also fitted a model without any regressors and added the EAP ability estimates from this

run. Table 2 shows the results of regressing these EAP estimates onto `level` and compares the results obtained with those obtained by ConQuest.

Table 2 OLS Regression Results Using Alternative Latent Ability Estimates

<i>Analysis Method</i>	<i>Latent Ability</i>		<i>Regression Coefficients</i>		<i>Conditional Variance</i>	<i>R²</i>
	<i>Mean</i>	<i>Variance</i>	<i>Constant</i>	<i>Level</i>		
Estimator*						
EAP with Regressor	0.49	1.08	0.335	0.277	1.065	0.02
EAP without Regressor	0.49	1.08	0.352	0.244	1.068	0.01
Direct ConQuest Estimation	0.49	1.23	0.338	0.275	1.214	0.02

* These are the ability estimates computed by ConQuest and then used as dependent variables in the analyses.

The mean is well estimated by the EAP latent ability estimates, but as in the previous sample analysis the variance is underestimated. The degree of underestimation is much less marked than it was in the previous sample analysis, but it is still noticeable. For the regression coefficients, we note that the EAP with regressor latent ability estimates are very close to the values produced by ConQuest. The EAP without regressor values are moderately biased, again due to their shrunken nature: the constant term is overestimated and the difference between the levels is underestimated. The conditional variances are again under-estimated by the EAP-based ability estimates.

LATENT MULTIPLE REGRESSION

The regressions undertaken in the last two sample analyses used a single regressor, `level`, which takes two values, 0 to indicate lower grade and 1 to indicate upper grade. This effectively meant that these two sample analyses were equivalent to two-sample *t*-tests. In ConQuest, up to 200 regression variables can be used simultaneously, and the regressors can be continuous numerical values. As a final sample analysis, we will show the results of analysing the data in `ex5.dat` using three regressors. The command file for this sample analysis (`ex5c.cqc`) is given in Figure 8. The only substantive difference between Figures 6 and 8 is in line 10, where the variables `gender` and `gbyl` are added.

```

1.  datafile ex5.dat;
2.  title  Australian TIMSS Mathematics Data--All Items;
3.  format gender 17 level 18 gbyl 19
      responses 20-176;
4.  labels << ex5.nam;
5.  codes 0,1,2,3,4,5,M;
6.  recode (R) (.);
7.  key
    13442341153114122133113244235223341323324535
    24322331211313511511324113143242223435112411
    14141222133113123244122143223342135422133251
    12253131111111111111111111111141!1;
8.  key
    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
    XXXXXXXX2X2XXXXX22XXXXXX2XX2XXXXXXXXXXXXXXXX2XXXX
    XX2X2XXX2XXXXX2XXXXX2XX2XXXXXXXXXX2XXXXXXXXXXXX2
    XXXXXXXX222222222222222222X2!2;
9.  key
    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
    XXXXXXXX3X3XXXXX33XXXXXX3XX3XXXXXXXXXXXXXXXX3XXXX
    XX3X3XXX3XXXXX3XXXXX3XX3XXXXXXXXXX3XXXXXXXXXXXX3
    XXXXXXXX333333333333333333X3!3;
10. regression level gender gbyl;
11. model item+item*step;
12. import init_parameters << ex5c.prm;
13. import init_reg_coefficients << ex5c.reg;
14. import init_covariance << ex5c.cov;
15. estimate !fit=no;
16. set p_nodes=1000;
17. show cases !estimates=eap >> ex5c.eap;
18. show !tables=3 >> ex5c.shw;

```

Figure 8 Sample Command File for a Latent Multiple Regression (All Items)

Figure 9 shows the contents of `ex5c.shw`, the population model parameter estimates for this third latent multiple regression sample analysis. The results reported in the figure show that the main effects of grade (`level`) and gender are 0.251 and -0.030, respectively, while the interaction between gender and grade (`gbyl`) is 0.052. The constant (0.351) is the estimated mean for male students in the lower grade. The estimated mean of female students in the lower grade is 0.321 ($=0.351-0.030$), of male students in the upper grade is 0.602 ($=0.351+0.251$), and of female students in the upper grade is 0.624 ($=0.351+0.251-0.030+0.052$).

NOTE: *The ConQuest population model is a regression model that assumes normality of the underlying latent variable, conditional upon the values of the regression variables. If you want to regress latent ability onto categorical variables or to specify interactions between variables, then appropriate contrasts and interaction terms must be created external to the ConQuest program. For instance, in this sample analysis, we have constructed the interaction between **gender** and **level** by constructing the new variable **gbyl**, which is the product of **gender** and **level**, and adding it to the data file.*

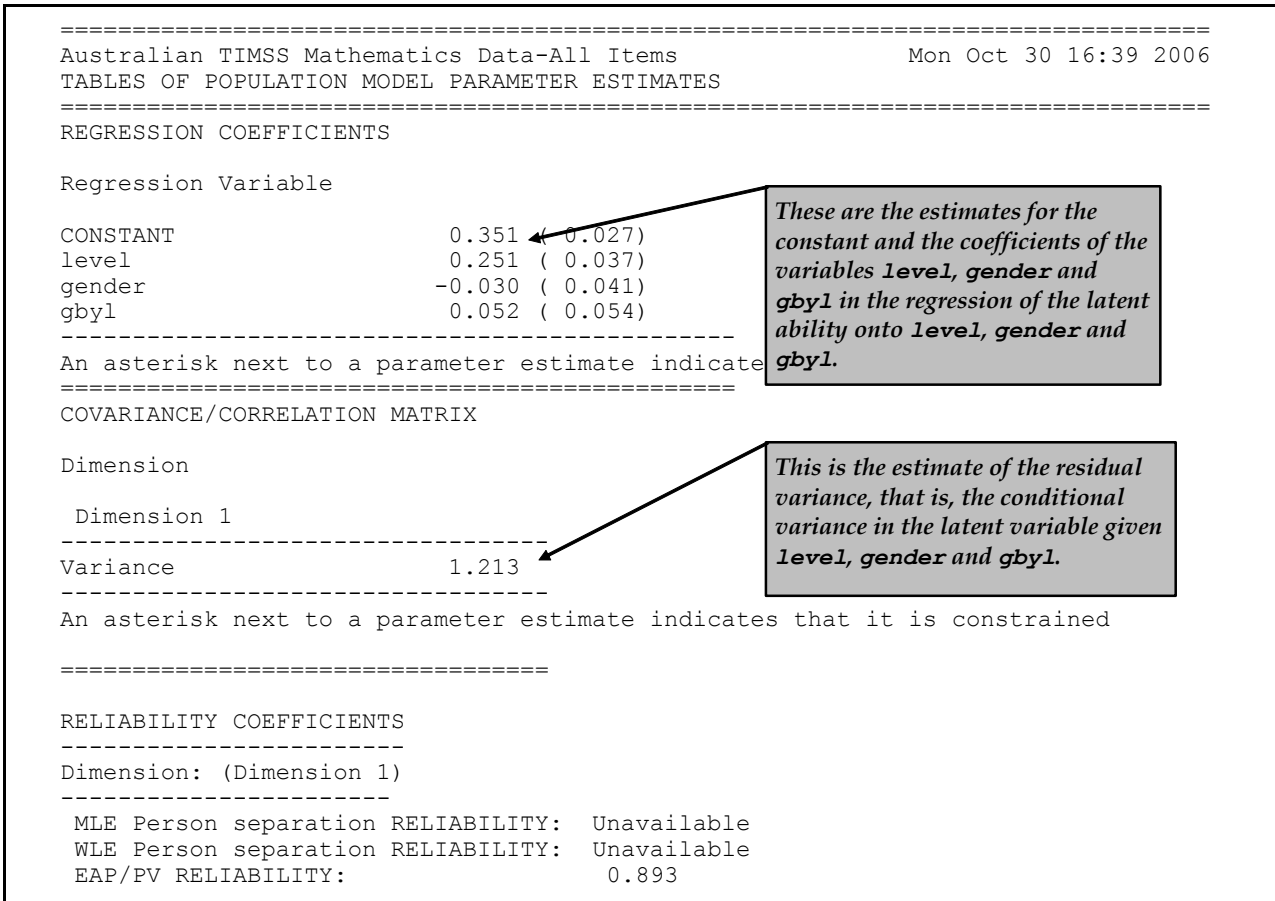


Figure 9 Population Model Parameter Estimates for the Latent Multiple Regression Sample Analysis

SUMMARY

In this tutorial, we have seen how to use ConQuest to fit unidimensional latent regression models. Our sample analyses have been concerned with using categorical regressors, but ConQuest can analyse up to 200 continuous or categorical regressors.

Some key points covered in this tutorial are:

- Secondary analyses using EAP and MLE ability estimates do not produce results that are equivalent to the 'correct' latent regression results. The errors that can be made in a secondary analysis of latent ability estimates are greater when measurement error is large.
- The key command can be used with a mixture of dichotomous and polytomous items.
- The show command can be used to create files of ability estimates. ConQuest provides both EAP and maximum likelihood ability estimates.
- The import command can be used to read files of initial values for parameter estimates.

REFERENCES

Adams, R. J., and Gonzalez, E. J. 1996. The TIMSS Test Design. In *Third International Mathematics and Science Study. Technical Report Volume 1: Design and Development*, M. O. Martin and D. L. Kelly (eds.), pp. 3-1-3-36 Center for the Study of Testing, Evaluation and Educational Policy. Boston, Massachusetts: Boston College.

- Adams, R. J., Wilson, M. R., and Wu, M. L. 1997. Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and Behavioural Statistics*, 22(1), 46-75.
- Beaton, A. L., Mullis, I. V. S., Martin, M. I., Gonzalez, E. J., Kelly, D. L., and Smith, T. 1996. *Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study*. Center for the Study of Testing, Evaluation and Educational Policy. Boston, Massachusetts: Boston College.
- Lokan, J., Ford, P., and Greenwood, L. 1996. *Maths & Science On the Line: Australian Junior Secondary Students' Performance in the Third International Mathematics and Science Study*. Melbourne: Australian Council for Educational Research.
- Lord, F. M. 1983. Unbiased estimators of ability parameters, of their variance, and parallel-forms reliability. *Psychometrika*, 48, 233-45.
- Lord, F. M. 1984. Maximum likelihood and Bayesian parameter estimation in item response theory. Research Report RR-84-30-ONR. Princeton, New Jersey: Educational Testing Service.
- Mislevy, R. J. 1984. Estimating latent distributions. *Psychometrika*, 49, 359-81.
- Mislevy, R. J. 1985. Estimation of latent group effects. *Journal of the American Statistical Association*, 80, 993-97
- Wright, B. D., and Stone, M. 1979. *Best Test Design: Rasch Measurement*. Chicago: MESA Press.
- Wu, M. L., Adams, R. J., Wilson, M. R., Haldane, S.A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software* [computer program]. Camberwell: Australian Council for Educational Research.