



UNDERSTANDING LABOUR FORCE SURVEY VARIABILITY FOR THE BASIN-BOUNDARY REGION

FALL, 2015



The Columbia Basin Rural Development Institute, at Selkirk College, is a regional research centre with a mandate to support informed decision-making by Columbia Basin-Boundary communities through the provision of information, applied research and related outreach and extension support. Visit www.cbrdi.ca for more information.

INTRODUCTION

The Columbia Basin Rural Development Institute (RDI) is mandated to report on conditions and trends relevant to the Columbia Basin-Boundary region. The data generated by Statistics Canada's (StatCan's) Labour Force Survey (LFS) is a critical component of RDI regional support. The RDI purchases annual data from Statistics Canada. This data is packaged specifically for the RDI and focuses on our region's context. The RDI further customizes the presentation of this data for more easily consumed information on our region's labour market. This information is disseminated to constituent organizations and private stakeholders throughout the region.

Labour market information is highly valuable to stakeholders in the public service, private enterprise and non-profit sectors. The LFS data helps form educational and labour policy, business plans, and strategic plans. Many stakeholders from a diverse spectrum of interests rely on LFS data for planning purposes.

There are a number of issues with the LFS accuracy when dealing with smaller population¹ regions. These impacts are amplified when we break-down results for our region by characteristics such as sector or nature of employment (i.e., part-time vs. full-time jobs). Again, when looking at rates and changes in values, impacts can be amplified by the comparison of two different estimates (ratios for rates, and temporal values for change).

This examination was motivated by observing wild swings in the LFS estimates of the rate of post-secondary education in the adult working-age population (25 – 54). Reported values, while seeming to mimic the provincial and national movement when smoothed over time, varied significantly from year to year.

Because of the heavy use this particular data-set sees, and concerns with reliability, the RDI has issued this methodology brief. This brief discusses the sources and impacts of LFS sampling variability (statistical accuracy).

LABOUR FORCE SURVEY

According to the Statistics Canada website, the Labour Force Survey (LFS) is:

“a monthly survey which measures the current state of the Canadian labour market and is used, among other things, to calculate the national, provincial, territorial and regional employment and unemployment rates. The survey results are used to make important decisions regarding job creation, education and training, retirement pensions and income support”.

SAMPLING

Households are randomly selected to represent the region in a series² of phone surveys. In British Columbia, 12,000 participants must be distributed to represent the province's eight Development Regions (DRs) to ensure reliable sub-provincial estimates. Roughly 1500 observations are taken per DR. For British Columbia LFS estimates, 1500 is also the “reliability threshold”, meaning that

¹ Within the rural context of the Basin-Boundary region, most communities would be considered “smaller

² Participants answer the LFS monthly, for six-months. This is referred to as a ‘rotation’.

estimates based on fewer observations are suppressed because they are deemed too unreliable (StatCan 2015).

ESTIMATION AND VARIABILITY

Statisticians take the answers from the sample and, based on this information, make estimations of how the whole population would answer the same questions (Wooldridge, 2002). This process, or estimation, can be very accurate under favourable conditions. With (relatively) small sample sizes, the likelihood of estimation error increases.

Using their knowledge of sampling errors in the LFS data, Statistics Canada publishes **Coefficients of Variation (CVs)** to express uncertainty in estimates (see Table 1 for a sub-set of Basin-Boundary CVs and resultant intervals of confidence). A CV can be thought of as a *standard error* that is expressed relative to the estimate itself. Standard error “is a measure that quantifies how different repeated samples might be from one another” (StatCan 2015, p.5). The concept dictates that, normally, 68% of samples can be expected to produce an estimate that is within one standard error of the actual value, and 95% of samples can be expected to produce an estimate that is within two standard errors of the actual value.

For instance, if the LFS estimates that 10,000 people have jobs in Trade with a CV of 5%, there is a 68% probability that the true population mean lies between 9,500 and 10,500 (10,000 plus or minus one CV, or 5%). There is a 95% probability that the true population value lies between 9,000 and 11,000 (10,000 plus or minus two CVs, or 10%).

Generally speaking, larger values (a bigger proportion of the whole population) are easier to accurately estimate than small values. Think about a room full of fifty senior citizens with two toddlers. If you wanted to estimate the number of people over fifty, a small random sample of ages would quickly determine that most in the room are over fifty. In fact, a random sample could easily miss one of the two toddlers. This would give the impression that all 52 people in the room were over fifty. This estimate would not be far from accurate (4% CV).

However, if we wanted to estimate the number of toddlers in the room, a random sample would have to be much bigger to ensure representation from the smaller population (toddlers). If our random sample missed both toddlers, we would mistakenly estimate the toddler population to be zero (100% CV).

REGIONAL CONTEXT

Three DRs are represented, in whole or in part, in the Basin-Boundary region. The Kootenay DR includes the Regional Districts of Kootenay Boundary, Central Kootenay, and East Kootenay. In our region, the Thompson-Okanagan DR is represented by the areas surrounding Golden and Revelstoke, and the Cariboo DR is represented by the area surrounding Valemount.

The smallest geography represented in the LFS is the Development Region. Given the different types of labour markets that are included in every single DR-level statistic, some of the statistics may over-represent some areas and under-represent others.

The areas around Golden, Revelstoke and Valemount are prone to this problem. When taken in the Basin-Boundary context, these areas are additionally challenged by the exclusion of most of their host DRs (and associated context) from the study area.

BASIN-BOUNDARY VARIATION

Using the methodology recommended by Statistics Canada in its *Guide to the Labour Force Survey*, we have analyzed the confidence intervals around the estimated values for a number of Basin-Boundary statistics. Table 1, below, gives confidence intervals for a sub-set of Basin-Boundary LFS indicators³.

Indicator Family	Indicator	CV	Interval for 68% Confidence
<i>Employment and Wages (Confidence interval is for employment only)</i>			
	Total	5.0%	+/- 3,500
	Goods-producing sector	7.5%	+/- 1,500
	Forestry, fishing, mining, quarrying, oil & gas	16.5%	+/- 1,155
	Construction	16.5%	+/- 1,155
	Manufacturing	16.5%	+/- 990
	Services-producing sector	5.0%	+/- 2,550
	Finance, insurance, real estate & leasing	33.3%	+/- 660
	Professional, scientific & technical services	25.0%	+/- 750
	Business, building & other support services	25.0%	+/- 750
	Educational services	20.0%	+/- 800
<i>Unemployment</i>			
	Unemployment Rate	19.0%	+/- 1%
<i>Labour Force Education</i>			
	<i>Labour Force Population w/</i>		
	Total, all education levels	5.0%	+/- 3,850
	High school graduate	10.0%	+/- 1,300
	Some post-secondary	5.0%	+/- 1,900
	Post-secondary certificate or diploma	7.5%	+/- 1,950
	University degree	10.0%	+/- 1,200
	<i>Percentage of Labour Force w/</i>		
	High school graduate	15.0%	+/- 2.5%
	Some post-secondary	10.0%	+/- 4.6%
	Post-secondary certificate or diploma	13.5%	+/- 4.6%
	University degree	15.0%	+/- 2.3%

Table 1: Select indicators, associated CVs and corresponding 68% confidence intervals

Source: StatCan, 2015; intervals calculated by the RDI

Consider ***Total Employment for the Kootenay DR (KDR)***, which was estimated at 67,500 in 2014. Table 1 gives a confidence interval of +/- 3,500 people. So, we can be 68% confident that the true population of employed persons in the KDR lies between 64,000 and 71,000 people. 67,500 is simply the mid-point of the best estimates.

Confidence intervals around percentages (such as the unemployment rate, or the rate of post-secondary attainment) apply the +/- percentage to the estimated rate. Included in Table 1 are a

³ For a more complete treatment of variability and confidence intervals, please refer to the CBRDI's white paper: *Interpreting the Labour Force Survey*

suite of indicators focused on education and the rate of educational attainment (the percentage of the population that has a particular level of education). Consider the Rate of Post-Secondary Attainment in the Kootenay DR.

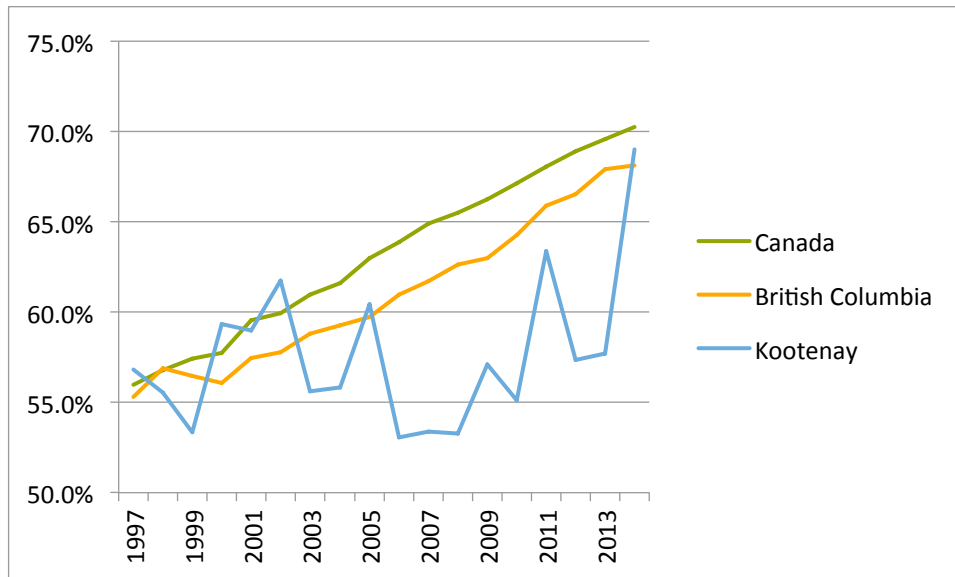


Figure 1: Rate of post-secondary education in the 25-54 year old labour force
Source: StatCan, LFS 2015a

Figure 1 shows how the estimates of Kootenay education trends have had a lot more year to year movement than the provincial and national indicators. Table 1 shows us that for estimates in the population ranges associated with post-secondary educational attainment⁴, we have CVs around 13.5%. For instance, the 2014 estimate is 69%. This tells us we can be 68% confident that the true value of Kootenay post-secondary educational attainment lies somewhere between 59.7% and 78.3% (69% +/- 13.5% of 69%).

IMPLICATIONS

Now that we've unpacked the LFS data accuracy, how do we deal with the situation? Fortunately, Statistics Canada has experience working with the most extreme remote rural living conditions in Canada; the Canadian territories. Through work with the northern territories, where sampling levels are extremely problematic, Statistics Canada has recommended using a three period moving average for the monthly estimates. This technique smooths out many of the unreliable swings in the data by allowing more people (different cohorts joining the survey every month) to participate, and by minimizing the impact of short-term changes in individual situations.

The RDI recommends adopting the three period moving average to our indicators that are prone to larger swings (Figure 2). Consider, again, the rate of post-secondary educational attainment in the Kootenay DR. Examination of the LFS estimates for the KDR (blue line in Figure 2) leaves one with the impression that Kootenay residents are wildly erratic in their relationship with education. The year-to-year changes defy explanation.

⁴ Note that Table 1 is based on the entire population, so the rates are lower than those presented in Figure 1, which is based only on the population between 25 and 54 years of age. While not comparable for analytic purposes, this is not a problem to illustrate the point.

However, if we look at the three year moving average (black line in Figure 2), we see far more understandable change tied to world events. There is still considerably more change than in the provincial and national estimates (orange and green lines, respectively), but it conforms to major events that may offer insight. For instance, there appears to be a sustained⁵ increase in the rate of education from 1999 to 2002. This corresponds to the “Dot.com” tech bubble bursting in 2000/2001. At that time, we saw a significant North American recession. We might expect to see people go back to school when the job market is tight.

Between 2002 and 2008, we see a sustained drop in the rate of education. This would correspond with the period of growth that North America enjoyed before the Great Recession of 2008/2009. While the economy is booming and the job market is hot, we may expect to see people forego education as the opportunity cost (lost wages in a booming job market) is too high. Then we see another sustained increase (roughly paralleling the national and provincial directions of movement) after 2008, when economic uncertainty and the down-sizing of the job market could have sent people back to school for retraining or upgrading while the job prospects are limited.

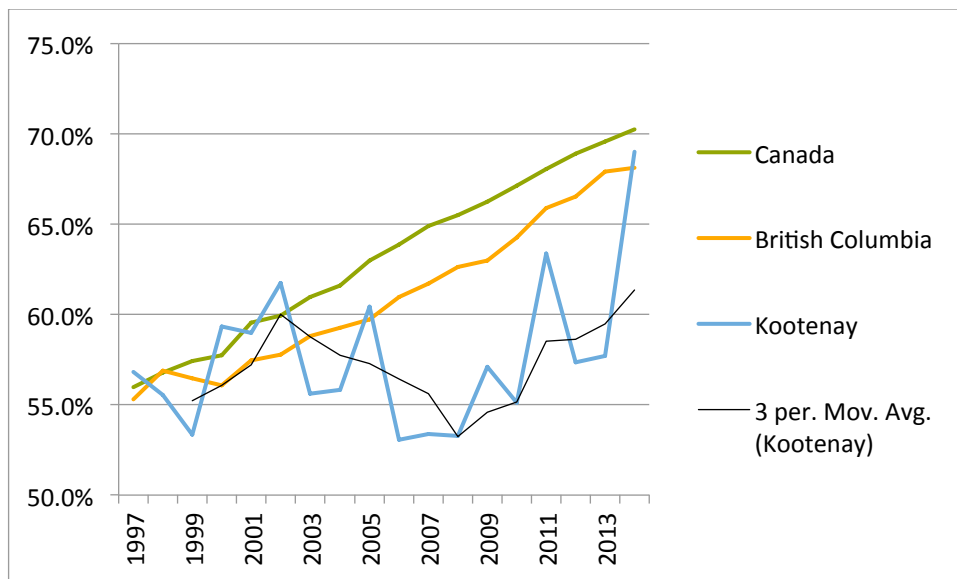


Figure 2: Rate of post-secondary education in the 25-54 year old labour force with 3 yr moving average
Source: StatCan, LFS 2015a

⁵ Since the three year average confirms it as a real trend.

RECOMMENDATIONS

The RDI intends to continue purchasing custom data sets from the LFS branch of StatCan. This data-set offers the only insight available into one of the most important influences on well-being: the labour market. While prone to sampling error, this data-set has interpretation issues, but it is still the only link that we have to timely quantitative data for our region.

The RDI will roll variation considerations into all of the analysis produced using the LFS. When using the LFS data as a basis for analysis, the RDI team intends to use both raw and three-year moving averages for comparison and consideration in the analysis. This step is especially important when looking at increasingly disaggregated indicators (those indicators that are so geographically and demographically specific as to exclude most of the population).

Changing our view to one that uses the three year moving average dampens the year to year variability that is associated with sample error and wide confidence intervals. This allows sustained trends to speak more loudly through the graphs while damping down sampling error variation. This comes at the cost of potentially damping down real swings in the data, but the cost of sampling error, in this set of circumstances, far outweighs the cost of missing the occasional swing.

Furthermore, in any discussions with potential user-groups (stakeholders with an interest in the labour market), we will recommend this approach and provide this report as background.

Finally, the RDI recommends that stakeholders use regional LFS estimates primarily to track trends (using the suggested three period moving averages) and focus less on the population estimates.

WORKS CITED

Columbia Basin Rural Development Institute. (2014). State of the Basin.

<http://www.cbrdi.ca/state-of-the-basin/digital-basin-portal/portal/>

Statistics Canada. (2015). *Guide to the Labour Force Survey*. Retrieved on August 15, 2015 from:

<http://www.statcan.gc.ca/pub/71-543-g/71-543-g2015001-eng.htm>.

Statistics Canada. (2015a). Labour Force Survey Custom Tables. Based on LFS Tables: 282-0002, 282-0004, 282-0008, and 282-0072

Wooldridge. (2002). *Introductory Econometrics: A Modern Approach*. South-Western College Pub.

APPENDIX A: INTERVAL OF VARIATION FOR LFS ESTIMATES

The Statistics Canada *Guide to the Labour Force Survey* offers a methodology for interpreting and applying the Coefficient of Variation tables (also provided in the guide).

1. Look up estimate (i.e. Labour Force in KDR in 2014 is 72,500).
2. Look up closest lower population size from Table 3, below (i.e. 36,300 is next value for BC – 77,200 is larger than our population, so we must look up next *lower* population).
3. Look up corresponding percentage(CV) from the column header (i.e. 5%).
4. Multiply the estimate by the CV (i.e. $72,500 \times 5\% = 3,625$)
5. Add and subtract this value to the original estimate (i.e. $72,500 \pm 3,625 = [68,875 \text{ to } 76,125]$)
6. Then, we are 68% confident that the true value lies somewhere between 68,875 and 76,125.
7. And, we are 95% confident that the true value lies somewhere between 65,250 and 79,750 (estimate $\pm (2 \times CV)$).

Coefficient of variation									
	1.00%	2.5%	5.00%	7.5%	10.00%	16.5%	20.00%	25.0%	33.30%
Canada	554.1	155.2	66.6	38	20.8	11.1	7.8	5.6	3.9
Newfoundland and Labrador	107.1	29.2	12	6.7	3.8	1.9	1.4	1	0.7
Prince Edward Island	28	8.3	3.8	2.2	1.2	0.7	0.5	0.4	0.3
Nova Scotia	98.7	29.3	13.3	7.8	4.3	2.4	1.7	1.2	0.9
New Brunswick	81.3	24	11	6.5	3.5	2	1.4	1	0.7
Quebec	365.5	109.6	51.6	30.8	16.2	9.4	6.5	4.7	3.5
Ontario	471.7	134.1	59.7	34.5	18.3	10.1	7	5	3.6
Manitoba	64.8	20.2	10.2	6.2	3.2	2	1.4	1	0.8
Saskatchewan	67.9	19.8	9.5	5.7	2.8	1.7	1.1	0.8	0.6
Alberta	223.3	68.8	34.2	20.9	10.6	6.5	4.4	3.3	2.5
British Columbia	262.4	77.2	36.3	21.5	11.1	6.4	4.4	3.2	2.4

Table 3: Coefficients of Variation for estimates of annual averages for Canada and the provinces

<http://www.statcan.gc.ca/pub/71-543-g/2015001/t005-eng.htm>

APPENDIX B: CVs FOR KOOTENAY DR ESTIMATES

Kootenay Development Region			
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV
<i>Employment and Wages</i>			
	Total	65-75	5.0%
	Goods-producing sector	17-22	7.5%
	Agriculture	1-3	>33%
	Forestry, fishing, mining, quarrying, oil & gas	5-8	16.5%
	Utilities	N/A	
	Construction	4-10	16.5%
	Manufacturing	4-7	16.5%
	Services-producing sector	45-56	5.0%
	Trade	10-15	10.0%
	Transportation & warehousing	2-4	20.0%
	Finance, insurance, real estate & leasing	1-3	33.0%
	Professional, scientific & technical services	1-4	25.0%
	Business, building & other support services	2-4	25.0%
	Educational services	2-6	20.0%
	Health care & social assistance	8-11	16.5%
	Information, culture & recreation	1-3	33.0%
	Accommodation & food services	5-8	16.5%
	Other services	2-4	25.0%
	Public administration	1-3	33.0%

Table 4: Coefficients of variation for Kootenay DR Employment and Wage estimates
Source: StatCan, 2015; intervals calculated by the RDI

Kootenay Development Region				
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV	Compound
Unemployment				
	Labour Force	72-81	2.5%	
	Unemployed	4-7	16.5%	
	Unemployment Rate		19.0%	= CV(LF) + CV (U)
Labour Force Education				
	<i>Labour Force Population w/</i>			
	Total (25-64)	58-67	5.0%	
	0 - 8 years	N/A	N/A	
	Some high school	4-7	16.5%	
	High school graduate	10-16	10.0%	
	Some post-secondary	7-11	16.5%	
	Post-secondary certificate or diploma	22-30	7.5%	
	University degree	9-13	10.0%	
	Bachelor's degree	4-10	16.5%	
	Above bachelor's degree	2-5	25.0%	
	<i>Percentage of Labour Force w/</i>			
	0 - 8 years		N/A	
	Some high school		21.5%	= CV(LF25-64) + CV (SHS)
	High school graduate		15.0%	= CV(LF) + CV (HSG)
	Some post-secondary		21.5%	= CV(LF) + CV (SPS)
	Post-secondary certificate or diploma		13.5%	= CV(LF) + CV (PSD)
	University degree		15.0%	= CV(LF) + CV (UD)
	Bachelor's degree		21.5%	= CV(LF) + CV (BD)
	Above bachelor's degree		30.0%	= CV(LF) + CV (GRAD)

Table 5: Coefficients of variation for Kootenay DR Unemployment and Education estimates

Source: StatCan, 2015; intervals calculated by the RDI

APPENDIX C: CVs FOR THOMPSON-OKANAGAN DR ESTIMATES

Thompson-Okanagan Development Region			
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV
<i>Employment and Wages</i>			
	Total	240-253	2.5%
	Goods-producing sector	55-65	5.0%
	Agriculture	5-9	16.5%
	Forestry, fishing, mining, quarrying, oil & gas	5-8	16.5%
	Utilities	1-3	33.0%
	Construction	24-30	7.5%
	Manufacturing	16-20	10.0%
	Services-producing sector	180-200	2.5%
	Trade	38-43	5.0%
	Transportation & warehousing	8-12	10.0%
	Finance, insurance, real estate & leasing	11-14	10.0%
	Professional, scientific & technical services	13-16	10.0%
	Business, building & other support services	7-11	16.5%
	Educational services	13-17	10.0%
	Health care & social assistance	30-34	7.5%
	Information, culture & recreation	11-13	10.0%
	Accommodation & food services	19-26	7.5%
	Other services	10-13	10.0%
	Public administration	8-11	16.5%

Table 6: Coefficients of variation for Thompson-Okanagan DR Employment and Wage estimates
Source: StatCan, 2015; intervals calculated by the RDI

Thompson-Okanagan Development Region				
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV	Compound
Unemployment				
	Labour Force	259-278	1.0%	
	Unemployed	11-24	10.0%	
	Unemployment Rate		11.0%	= CV(LF) + CV (U)
Labour Force Education				
	<i>Labour Force Population w/</i>			
	Total (25-64)	201-224	2.5%	
	0 - 8 years	1-3	33.0%	
	Some high school	12-20	10.0%	
	High school graduate	44-50	5.0%	
	Some post-secondary	16-23	10.0%	
	Post-secondary certificate or diploma	76-91	2.5%	
	University degree	34-45	5.0%	
	Bachelor's degree	23-31	7.5%	
	Above bachelor's degree	11-15	10.0%	
	<i>Percentage of Labour Force w/</i>			
	0 - 8 years		35.5%	
	Some high school		12.5%	= CV(LF25-64) + CV (SHS)
	High school graduate		7.5%	= CV(LF) + CV (HSG)
	Some post-secondary		12.5%	= CV(LF) + CV (SPS)
	Post-secondary certificate or diploma		5.0%	= CV(LF) + CV (PSD)
	University degree		7.5%	= CV(LF) + CV (UD)
	Bachelor's degree		10.0%	= CV(LF) + CV (BD)
	Above bachelor's degree		12.5%	= CV(LF) + CV (GRAD)

Table 7: Coefficients of variation for Thompson-Okanagan DR Unemployment and Education estimates
Source: StatCan, 2015; intervals calculated by the RDI

APPENDIX D: CVs FOR CARIBOO DR ESTIMATES

Cariboo Development Region			
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV
<i>Employment and Wages</i>			
	Total	75-85	2.5%
	Goods-producing sector	21-26	7.5%
	Agriculture	1-2	>33%
	Forestry, fishing, mining, quarrying, oil & gas	4-7	20.0%
	Utilities	N/A	
	Construction	6-8	16.5%
	Manufacturing	8-11	16.5%
	Services-producing sector	54-62	5.0%
	Trade	11-14	10.0%
	Transportation & warehousing	4-6	20.0%
	Finance, insurance, real estate & leasing	2-3	33.0%
	Professional, scientific & technical services	3-4	25.0%
	Business, building & other support services	2-3	33.0%
	Educational services	4-7	20.0%
	Health care & social assistance	9-11	16.5%
	Information, culture & recreation	2-3	33.0%
	Accommodation & food services	5-8	16.5%
	Other services	3-5	25.0%
	Public administration	2-4	33.0%

Table 8: Coefficients of variation for Cariboo DR Employment and Wage estimates
Source: StatCan, 2015; intervals calculated by the RDI

Cariboo Development Region				
Indicator Family	Indicator	Range of LFS Estimates (000s)	CV	Compound
Unemployment				
	Labour Force	85-92	2.5%	
	Unemployed	4-11	16.5%	
	Unemployment Rate		19.0%	= CV(LF) + CV (U)
Labour Force Education				
	<i>Labour Force Population w/</i>			
	Total (25-64)	68-76	5.0%	
	0 - 8 years	<2	>33%	
	Some high school	5-9	16.5%	
	High school graduate	16-20	10.0%	
	Some post-secondary	4-7	16.5%	
	Post-secondary certificate or diploma	25-31	7.5%	
	University degree	9-14	10.0%	
	Bachelor's degree	7-9	16.5%	
	Above bachelor's degree	2-6	20.0%	
	<i>Percentage of Labour Force w/</i>			
	0 - 8 years		>38%	
	Some high school		21.5%	= CV(LF25-64) + CV (SHS)
	High school graduate		15.0%	= CV(LF) + CV (HSG)
	Some post-secondary		21.5%	= CV(LF) + CV (SPS)
	Post-secondary certificate or diploma		13.5%	= CV(LF) + CV (PSD)
	University degree		15.0%	= CV(LF) + CV (UD)
	Bachelor's degree		21.5%	= CV(LF) + CV (BD)
	Above bachelor's degree		25.0%	= CV(LF) + CV (GRAD)

Table 9: Coefficients of variation for Cariboo DR Unemployment and Education estimates

Source: StatCan, 2015; intervals calculated by the RDI