

OREGON LITERACY PROJECT

Project Description

The Oregon Progress Board commissioned a statewide survey of adult functional literacy proficiency that was conducted in 1990 and reported shortly thereafter (Oregon Progress Board, 1991). A random sample of about 2,000 Oregon adults, age 16-65, were interviewed and assessed in their homes. Results from this Oregon Literacy Survey, which included a hands-on standardized assessment of functional literacy developed by the Educational Testing Service (ETS), were reported on three scales of literacy proficiency: Prose, Document and Quantitative literacy.

Each person's proficiency was estimated separately for Prose, Document and Quantitative literacy on a 0-500 point scale. The Oregon Progress Board established three proficiency benchmarks for each scale: *Basic* (scale scores of 225 and above), *Intermediate* (scale scores of 325 and above), and *Master* (scale scores of 375 and above). The Intermediate proficiency level was adopted by the Progress Board as Benchmark #27 of the State's strategic vision, *Oregon Shines*.¹

Table 1 displays the percentage of Oregon adults (age 16-65) who met the Intermediate (325 and above) benchmark in the 1990 survey.² Across the three literacy scales, 36 to 41% of Oregon adults age 16-65 met or exceeded the benchmark in 1990.

	Prose	Document	Quantitative
Percent At or Above Literacy Benchmark	41.1	35.5	39.0

Table 1. Percent of Oregon adults, age 16-65, meeting or exceeding literacy benchmarks. Source: Oregon Literacy Survey.

In its periodic reports to the Legislative Assembly, the Oregon Progress Board attempts to update each benchmark indicator with the most recent data available in order to track progress towards meeting state strategic goals. For most of the benchmarks, it is possible to update progress from recent data collected by local, state or national agencies or research studies. For the adult literacy benchmark, however, no statewide data have been collected since the original Oregon Literacy Survey data of 1990. As a result, some thirteen years later, the Progress Board has understandably been limited to grading the

¹ It is important to note that Oregon's 325 benchmark is substantially higher than the national benchmark of 275 later adopted by the National Educational Goals Panel (1997) for the National Adult Literacy Survey (NALS) that used the same assessment methodology (Kirsch et al, 1993). The national standard for adult literacy is Level 3 or above (out of five levels within the 0-500 scales), whereas Oregon's benchmark of 325 and above is Level 4 and above.

² As a check, these results were calculated independently from the original Oregon Literacy survey data and agree exactly with the original figures reported by the Oregon Progress Board (1991, Table 1).

state's progress on adult literacy skills as "Unknown" (Oregon Progress Board, 2003, p. 24).

Although a follow-up national assessment of adult literacy (the National Assessment of Adult Literacy or NAAL) is underway that will provide literacy assessments for 2003 that are similar to those conducted by Oregon in 1990, the national sampling design is not expected to permit reliable and valid statewide estimates of Oregon's adult literacy proficiencies. Although all the states including Oregon had the opportunity to contract for an enhanced statewide survey as part of the NAAL, Oregon did not contract for the additional statewide surveying.³ Had Oregon funded such a statewide survey, the data could be used to provide valid and reliable estimates of the adult literacy benchmark for 2003 (although these data would likely not be available until mid-2005).

The Oregon Progress Board thus needs other means for obtaining updated estimates of its adult literacy benchmarks. The current project is intended to produce new estimates through three different approaches. Two approaches are based on statistical modeling of relationships between adult literacy surveys and decennial Censuses. These two approaches will attempt to provide reasonable estimates of the benchmarks for 2000 without the expense of collecting new assessment data. The third approach will propose procedures (but not implement them) to add an assessment component to the biannual Oregon Population Survey as a possible relatively inexpensive means of obtaining updated adult literacy benchmarks when needed in the future.⁴

This project's first two tasks use two different statistical modeling techniques to estimate Oregon's adult literacy benchmarks for the year 2000. Each modeling method is based on previously developed and published techniques. The first method, synthetic estimation (Reder, 1996), has been used to develop estimates of adult literacy proficiencies for counties and states based on key Census indicators that are correlated with direct measures of adult literacy in state and national surveys. Such models have been used to extrapolate literacy results across space (i.e., to different geographical areas than those sampled in surveys) at a fixed point in time.

The project's second task builds a population model of adult literacy proficiency and then estimates changes in the population's literacy proficiency over a period of ten years (i.e., between 1990 and 2000) based on assumptions about how population changes are reflected in changes in literacy levels, considering the skills of individuals who die or leave the state over the ten year period, mature from children into young adults over that period, gain additional education over that period, and so forth. These modeling techniques, developed by Reder and Edmonston (2000) for predicting changes in national literacy levels over a decade, will be applied to predicting how Oregon's adult literacy will change over the decade between 1990 and 2000.

³ The cost to a state was \$725,000 to include an additional state-valid sample of 1,000 interviews in the NAAL. Only six states elected to contract for these state assessments: Kentucky, Maryland, Massachusetts, Missouri, New York and Oklahoma (www.nces.ed.gov/naal/state/costs.asp).

⁴ The Oregon Population Survey is conducted every two years and is next scheduled for 2004.

Neither of these previously developed models can be immediately applied to Oregon's adult literacy data. Each must be adapted to differences between the national data to which they were originally applied and the Oregon data. In the case of the synthetic estimates (Task 1), the model is treading over new statistical modeling terrain. A number of techniques described in Reder (1996) were used to assess the validity and reliability of such synthetic estimates at a given point in time. In the current project, an attempt will be made to apply models validated for estimating adult literacy proficiencies at one point in time (with respect to the 1990 Census, i.e.) to Census data collected in a subsequent decade (in the 2000 Census, i.e.). In the case of the estimates produced by a population model of Oregon (Task 2), new modeling challenges are involved as well in moving from the national to the state level. The impact of state-to-state migration, for example, is a variable that did not need to be considered in modeling national change, whereas it may have considerable influence on modeling changes in the state of Oregon's population. The validity of estimates produced by either type of model probably cannot be determined until new direct assessment data become available.⁵ To the extent that the two models appear free of major technical problems and produce comparable estimates for Oregon's adult literacy benchmarks, we can be relatively confident in the reasonableness of the estimated benchmark levels. If results from both approaches show moderate and consistent changes in Oregon's adult literacy levels – whether up or down – this would provide strong evidence of improvements (or decreases) in literacy levels between 1990 and 2000.

Task 3 of the project explores the technical issues, design considerations and costs of several alternative means of conducting a statewide adult literacy assessment. Options to add a literacy assessment component on to existing surveys are considered, along with options to conduct independent adult literacy surveys. By capitalizing on the planning, sampling and background interviewing that are already done as part of an existing survey, an add-on approach to assessing adult literacy proficiency may provide relatively cost-effective estimates for the adult literacy benchmark. Even if the estimates provided by Tasks 1 and 2 prove to be satisfactory, it may be worthwhile to build such an add-on component to future Oregon Population Surveys.

Task 1

General approach. Task 1 produces county and statewide estimates of the adult literacy benchmarks for Oregon, using statistical modeling techniques developed by Reder (1996). In the original work, regression models for adult literacy proficiency were developed that predicted literacy measured at the county and state levels in the 1992 National Adult Literacy Survey (NALS). The predictive models used background variables such as educational attainment and English language proficiency (among non-

⁵ Such validation may be possible, for example, using the new national NAAL literacy data that are expected to become available for secondary analysis in mid-2005.

native speakers) that were measured in the same way in the NALS and in the long-form of the 1990 U.S. Census. Technical descriptions of these models and the techniques used to validate them are described elsewhere (Reder, 1996). Once satisfactory models were developed, they were applied to the 1990 Census data from geographical areas that were not (well-) sampled in the NALS to produce synthetic estimates of adult literacy proficiency.

In Task 1, we try to apply such modeling techniques – originally designed to link the 1992 NALS and 1990 Census – to link the 1992 NALS and the 2000 Census data in order to produce more current estimates of the adult literacy benchmarks for Oregon and its counties. We face several challenges in attempting to use the same modeling techniques to construct the new synthetic estimates:

- ? Relationships between Census variables and literacy proficiencies that were validated for 1990-1992 data may not apply to linking the 2000-1992 data. Models that are well-specified for the 1990 data may not prove to be well-specified for the 2000 data.
- ? The NALS survey and the synthetic estimates derived from the 1990 Census dealt with the population of adults age 16 and above. The Oregon adult literacy benchmarks (emerging from the 1991 Oregon Literacy Survey) deal only with adults age 16-65.
- ? The benchmarks estimated by Reder (1996) were for proficiency Levels 1 and 2 (i.e., the percentage of adults scoring below 225 or below 275), whereas Oregon's benchmark is the percentage of adults scoring above 325 (Level 4 or 5).
- ? Only a subset of the Census variables used in modeling the 1990 data is available for modeling the 2000 data. Some variables have been changed by Census (e.g., occupational categories), others are not reported in standard tables in the same way in 2000 as in 1990 (e.g., work-preventing disability), and others are unlikely to retain a fixed relationship to literacy (e.g., labor market variables) between 1990 and 2000.

The caveat here is that because of these many changes, the validity and precision established for the synthetic estimates produced in the original modeling cannot be assumed to apply to the current efforts. They will have to be independently established.

Variables used. The first step in the process is to develop predictive models linking Census variables to the 1992 NALS national literacy data. In developing these models, only the adult population age 16-65 is considered since Oregon's literacy benchmarks are defined for that subpopulation. Furthermore, only a subset of the variables used by Reder (1996) that are common to the 1990 and 2000 Census variables are used so that the resulting models can be used to produce estimates for both 1990 for comparison with

directly obsessed skills) and for 2000 Census data for estimating current benchmark levels in Oregon.

As detailed below, a small set of variables – common to the NALS survey, the 1990 Census and the 2000 Census – were used in constructing these models. These variables contain the proportion of adults (age 16-65) in a geographic area such as a county who are Black, who are Hispanic, who are recent immigrants, who have various levels of education, and who have various levels of English language proficiency. There is also an indicator variable showing which of four major regions of the United States the given geographical area is in.

Educational attainment. Similar questions were asked in NALS and the long form of the 1990 and 2000 Censuses. Answers to these questions were recoded into a common seven-category variable reflecting the highest level of education:

1. less than high school
2. some high school but no diploma or equivalency
3. high school diploma or equivalency
4. some college but no degree
5. two-year college degree
6. four-year college degree
7. advanced degree

The proportion of a county or state population falling into each of these categories was calculated for the population 18-64 (since that is the closest age-range available in the 2000 Census SF3 tables). The regression analyses used six of these seven proportions as predictors, omitting the proportion at the lowest level as a reference category.

English language proficiency. Both the NALS and the long-form Census asked how well individuals spoke English (on a four-point scale) only if a response to previous question indicated a language other than (or in addition to) English was spoken in the home. The common four-point scale was:

Very well – Well -- Not well -- Not at all

For our analyses, individuals living in households in which only English was spoken were deemed to speak English “very well”. The lowest two categories (“not well” and “not at all”) were combined in our analyses. The regression models used two variables, corresponding to the proportion of a population (age 18-64, the age range tabulated in 2000 Census SF3) in the “Very well” and “Well”, categories, omitting the (“Not well” or “Not at all”) as a reference category.

Recency of immigration. Individuals born in the U.S. or who had immigrated to the U.S. at least five years before the NALS or Census were coded as “0” while individuals who had immigrated to the U.S. with five years of the NALS or Census were coded “1”. The regressions used the proportion of the population coded “1” (i.e., who were recent

immigrants) as a predictor variable. This proportion was calculated directly for individuals age 16-65 in NALS and the 1990 PUMS. It is tabulated only for all individuals (regardless of age) in 2000 Census SF3. Comparison of the age-restricted and age-unrestricted estimates in the 1990 PUMS indicates that this is a reasonable estimate.

Black. This variable was calculated as the proportion of the target population identifying as African-American. In both the NALS and 1990 Census PUMS data, this was estimated directly for the population within the target 16-65 age range. For the 2000 Census SF data, this proportion was estimated from the SF1 tables for ages 15-64 rather than 16-65 because of the age “buckets” available in the SF1 tables. Individuals who identified more than one race in the 2000 Census were excluded from this calculation.

Hispanic. This variable was calculated as the proportion of the target population identifying as Hispanic or Latino. In both the NALS and 1990 Census PUMS data, this was estimated directly for the population within the target 16-65 age range. For the 2000 Census SF data, this proportion was estimated from the SF1 tables for ages 15-64 rather than 16-65 because of the age “buckets” available in the SF1 tables.

Region. Four dummy variables were constructed for the NALS and Census data sets to indicate the Census region containing the given county or state. The West region (containing Oregon) was the reference category.

Regression models. Counties (occasionally clusters of adjacent counties) were the primary sampling unit of the NALS. County subsamples are representative (albeit often with small subsample sizes) of the counties from which they were drawn. The mean value of the predictor variables and the benchmark indicators (the percentage of adults age 16-65 in the county with Prose, Document and Quantitative proficiencies at or above scale scores of 325) were calculated for each of these aggregates. Separate weighted least squares⁶ regression models were estimated for each of the three benchmark proportions.⁷

Table 2 summarizes the fit of the regression models for the three benchmarks. Details of the regression models are tabled in Appendix A. Results for the three models are about the same. Each shows a reasonably satisfactory fit with the national sample of counties within the NALS data, predicting 68-71 percent of the variance in the dependent variables (i.e., the benchmark proportions). The techniques described in Reder (1996) for examining and validating the fit of these models were used again here, and indicate that the fit of the data is generally satisfactory although not as close as that obtained in the earlier modeling. This is probably to be expected given the limited age range (16-65 rather than 16 and above), the limited subset of predictor variables available, and the higher level of the benchmark being predicted here (all of which likely decrease the

⁶ The weight used for a given county aggregate was $N/(p*(1-p))$, where N is the size of the county subsample and p is the observed value of the dependent variable (benchmark proportion) for the county.

⁷ Preliminary analyses indicated that the better-fitting models were obtained with weighted least squares linear regressions than with logit or logistic regressions.

ability of the model to predict the benchmark proportions in these 1992 NALS data. Nevertheless, the fit of the models appears to be generally satisfactory.

	Prose	Document	Quantitative
N	391	388	391
R	.847	.833	.845
Adjusted R ²	.707	.683	.704
Mean residual	.018	.019	.016

Table 2. Summary of regression models. See text for details.

Predicting state benchmark levels. The generally satisfactory fit of the models can be seen in Figures 1, 2 and 3 for the Prose, Document and Quantitative benchmarks, respectively. Each figure shows the observed and predicted benchmark values for the eleven states in the 1992 NALS assessment that had large statewide samples. There is a pair of bars shown for each state. The left bar in each pair shows the Oregon benchmark value observed for the given state, whereas the right bar of the pair shows the corresponding benchmark value predicted by applying the regression equation to the state's 1990 Census variables. As can be seen, the model does a reasonably good job predicting statewide benchmarks from 1990 Census data.

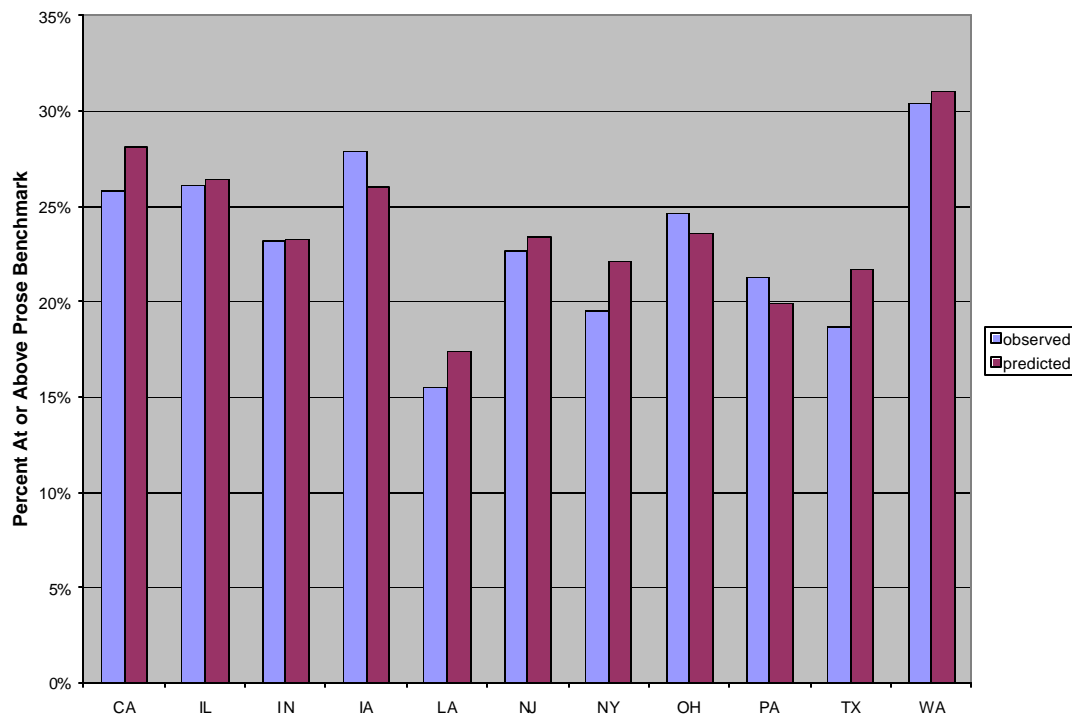


Figure 1. Observed and predicted Prose literacy benchmark values for eleven states.

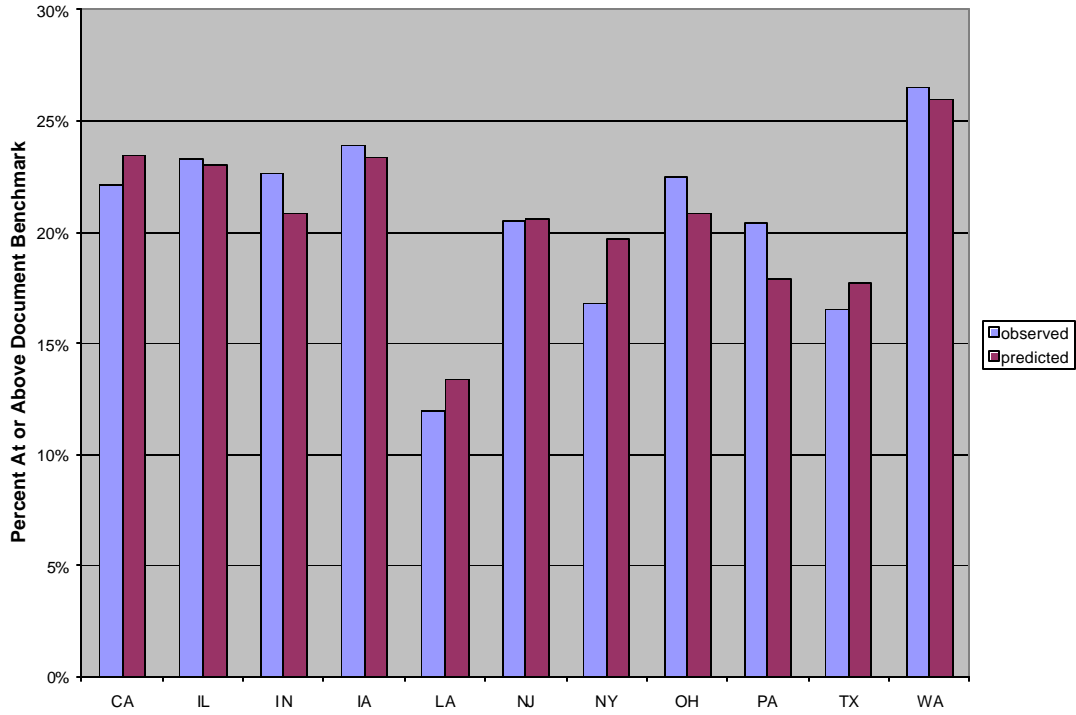


Figure 2. Observed and predicted Document literacy benchmark values for eleven states.

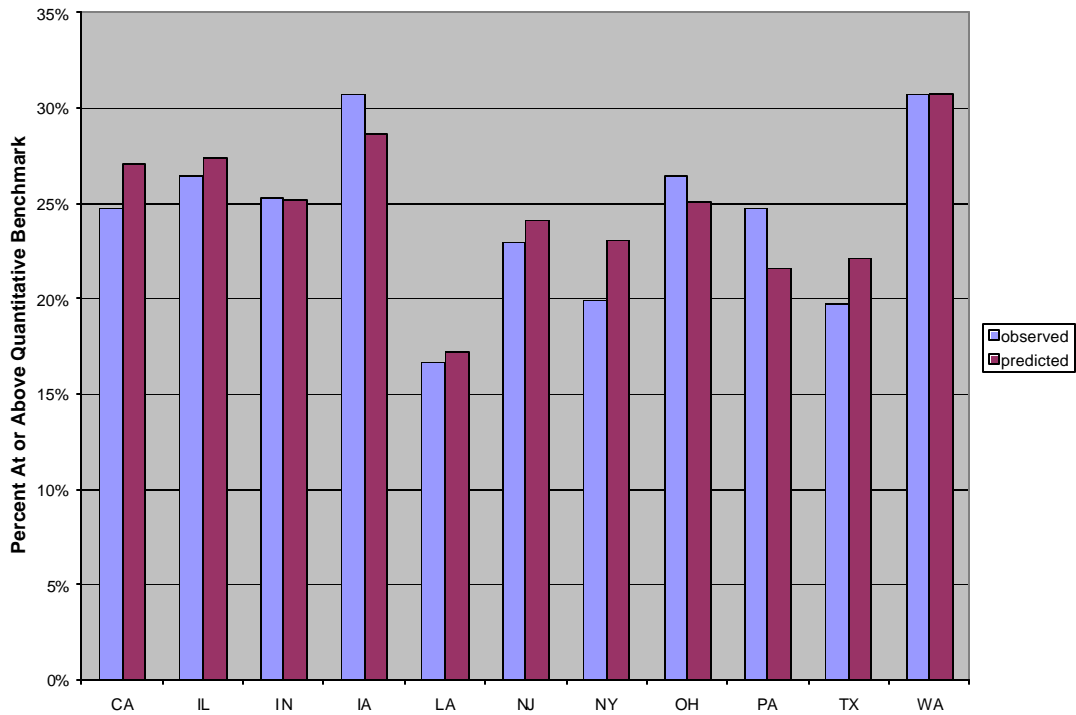


Figure 3. Observed and predicted Quantitative literacy benchmark values for eleven states.

Predicting benchmark levels for Oregon. Although the regression models perform satisfactorily predicting the observed benchmark levels for the eleven states from their 1990 Census data (as seen above in Figures 1, 2 & 3), they do not accurately predict the benchmark values observed in the 1990 Oregon Literacy Survey from Oregon's 1990 Census data. Figures 4, 5 and 6 are parallel to Figures 1, 2 and 3, respectively, showing again the data for the eleven states along with a new pair of bars for Oregon on the right-hand side. The left hand bar of the Oregon pair is the benchmark statistic for the state measured in the 1990 Oregon Literacy Survey (i.e., the data from Table 1), and the right hand bar is the predicted value of the regression model applied to Oregon's 1990 Census data.

As Figures 4, 5 and 6 show, the benchmark values predicted by the models are *substantially* lower than the values measured in the Oregon Literacy Survey. Oregon has much higher observed benchmark levels than any of the comparator states,⁸ even its neighboring state of Washington, which has an adult population quite similar to Oregon (and for which the model predicts similar adult literacy proficiencies). Since the benchmarks being developed here for the year 2000 are to be generated by applying these same models to Oregon's 2000 Census data, such a discrepancy is quite problematic and requires further analysis. There are a number of possible explanations for the discrepancy. One possibility is that Oregon's adult literacy proficiencies differ sharply from those of other states (i.e., are much higher) in ways that cannot be accounted for by the Census-based predictor variables. Another possibility is that even though the literacy assessment technology in the NALS national survey and the Oregon Literacy Survey were very similar (and designed to be directly comparable), some systematic difference between the sampling and/or assessment procedures used in the two studies made their results highly discrepant.

Exploring the distinctiveness of adult literacy proficiencies in Oregon. To explore these possibilities more closely, we need to look more closely at the differences between the Oregon and U.S. literacy survey results. Figure 7 compares the adult literacy benchmark values for Oregon's 16-65 year old adult population as measured by the Oregon Literacy Survey in 1990 and the corresponding benchmark values measured by the NALS in 1992 for the national population of adults age 16-65. For each literacy scale, the percentage of 16-65 year olds meeting the benchmark is much higher in Oregon than in the nation as a whole. Certainly some of these differences may be attributable to Oregon having a much less diverse population than the country as a whole since immigrant and minority populations as a whole have lower (English) literacy proficiencies. Figure 8 therefore compares Oregon and U.S. adults age 16-65 who are native-born and not members of minority group (i.e., U.S.-born, white and not Hispanic). Although the benchmark levels are somewhat higher for this subpopulation in both the Oregon and US data, the Oregon benchmark levels remain substantially higher than their

⁸ It is important to note that the models are not based just on fitting the eleven states with large state samples with the NALS study; the models fit a large and representative set of counties from across the U.S.

U.S. counterparts. The large difference thus cannot be explained in terms of different demographics of the U.S. versus Oregon.

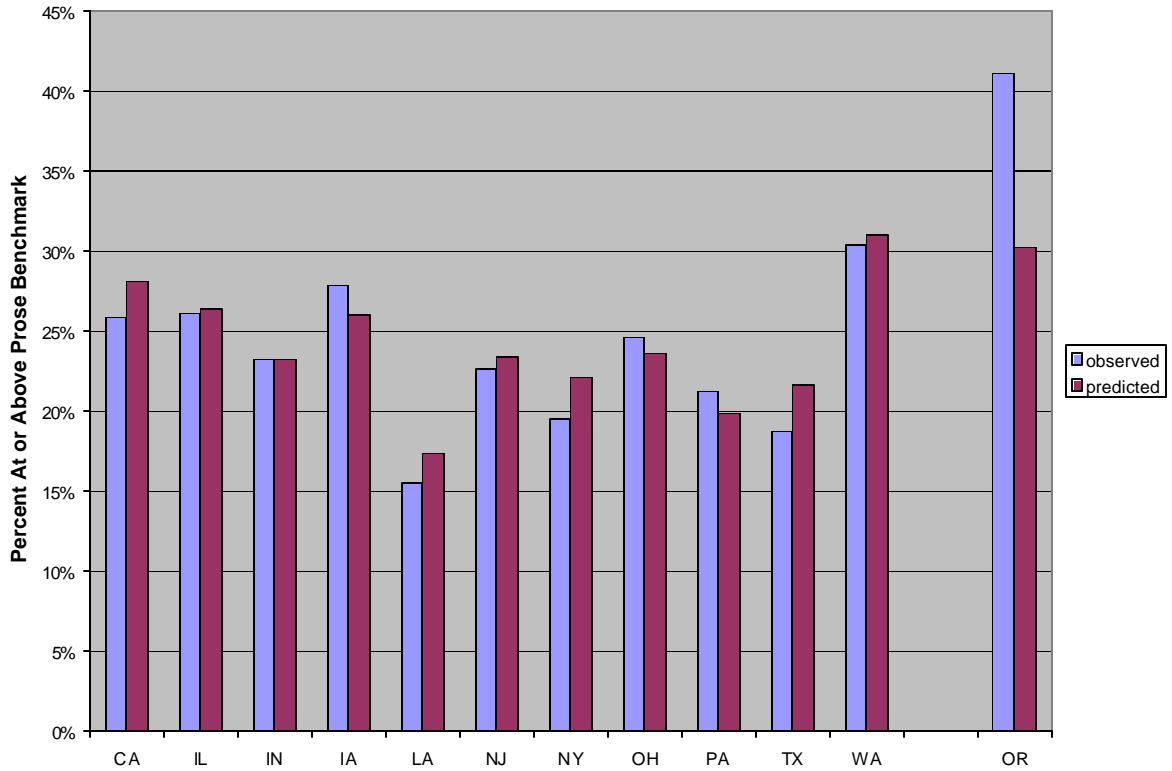


Figure 4. Observed and predicted Prose literacy benchmark values for eleven states and Oregon.

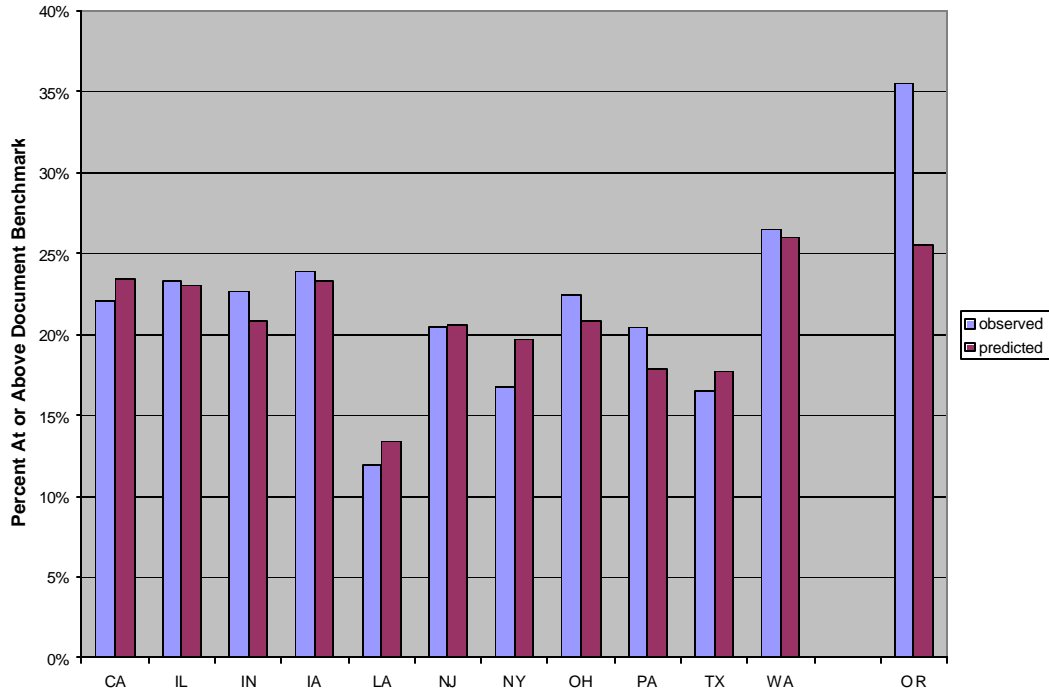


Figure 5. Observed and predicted Document literacy benchmark values for eleven states and Oregon.

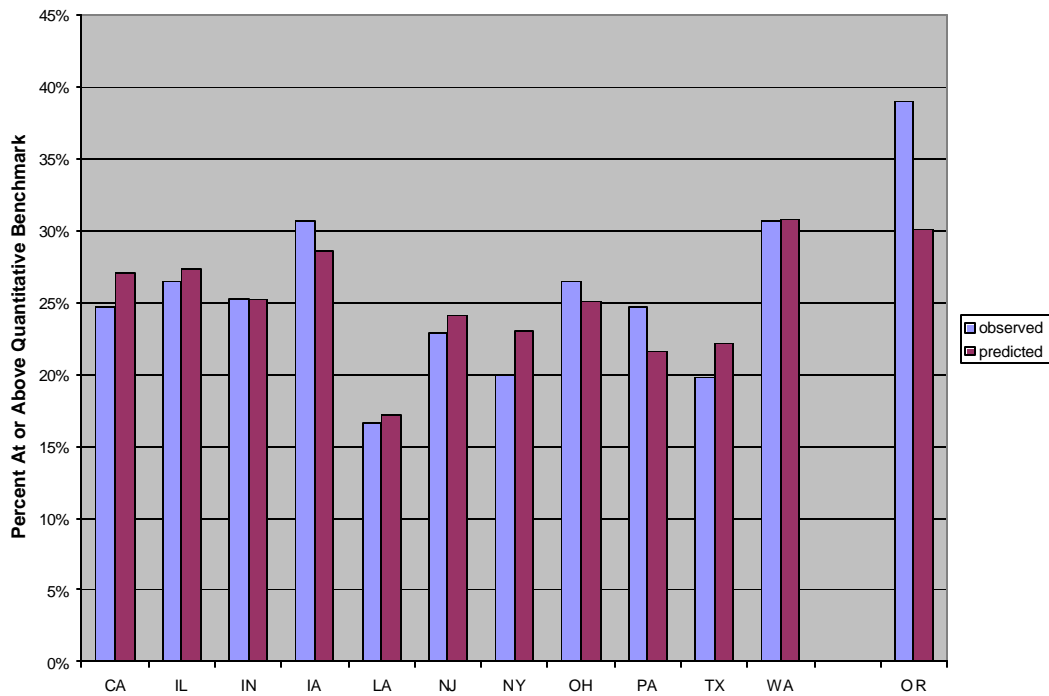


Figure 6. Observed and predicted Quantitative literacy benchmark values for eleven states and Oregon.

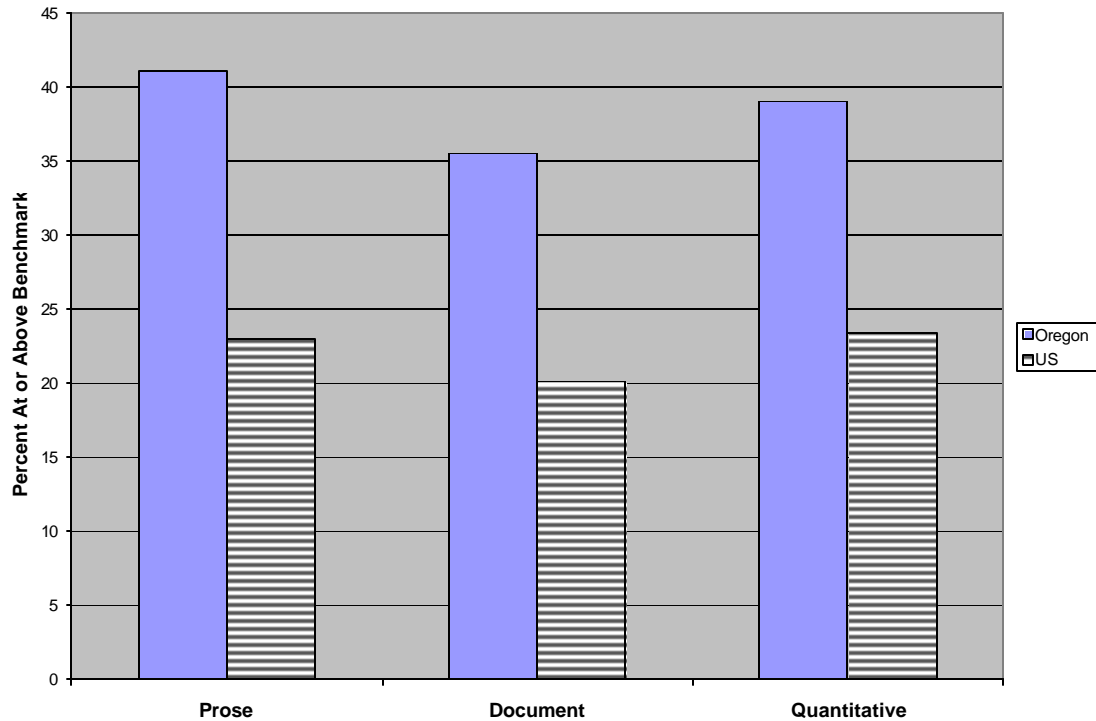


Figure 7. Literacy benchmark values for adults age 16-65. Source: Author calculations from the 1990 Oregon Literacy Survey (Oregon data) and the 1992 National Adult Literacy Survey (U.S. data).

Since educational attainment and adult literacy proficiency are very closely connected (Kaestle et al, 2001; Reder, 1998), it could be that different distributions of educational attainment in Oregon’s 16-65 year-old population underlie the observed differences in overall adult literacy proficiency. To examine this possibility, Figure 9 plots the assessed adult literacy proficiencies (Prose scale) for the Oregon and US populations (age 16-65, US-born, white & not Hispanic) as a function of educational attainment. The corresponding figures for the Document and Quantitative scales look very similar and are shown in Appendix B. As can be seen in Figure 9, benchmark levels in Oregon are substantially higher than in the US even when race, ethnicity, nativity and education are comparable.

Thus even taking demographic and educational differences into account, the assessed adult literacy levels are much higher in Oregon than elsewhere in the U.S. Another possibility we need to consider is that even though the NALS and Oregon Literacy Survey shared a common assessment methodology and were designed to be comparable, unintended methodological differences in the implementation of the surveys and assessment procedures resulted in the higher literacy levels in Oregon.

Fortunately, we can examine this possibility. The NALS survey itself contained a small number of respondents from Oregon (as part of the nationally representative but not

necessarily state-representative sample). There were 70 such Oregon NALS respondents in the 16-65 age range. Their benchmark levels were 48%, 39% and 38%, respectively, for Prose, Document and Quantitative literacy. This Oregon subsample of NALS is not statistically representative of the Oregon statewide population, but is representative of the counties from which this Oregon subsample was drawn for NALS: the tri-county area of Multnomah, Clackamas and Washington counties. Since the Oregon Literacy Survey database identifies the zip code area in which each respondent lived, it is possible to estimate the literacy benchmarks for the same tri-county area. When this is done, the Oregon Literacy Survey estimates for benchmarks on the Prose, Document and Quantitative scales for the tri-county area are 44%, 41%, and 44%, respectively. Table 3 displays these results from the tri-county area from the NALS and the Oregon Literacy Survey. The subsample size in NALS is fairly small (n=70), the subsample size in the Oregon Literacy Survey (n=768) is substantial.

Benchmark	NALS (n=70)	Oregon Literacy Survey (n=768)
Prose	48.4%	43.5%
Document	39.2%	40.9%
Quantitative	38.4%	43.7%
<i>(Average)</i>	42.0%	42.7%

Table 3. Comparison of Oregon benchmarks for the same *tri-county area* of Oregon, as estimated by the National Adult Literacy Survey of 1992 and the Oregon Literacy Survey of 1990.

Overall, the two surveys estimate relatively similar values for the Oregon benchmark levels. For one of the scales, the NALS estimate is somewhat *higher* than that of the Oregon Literacy Survey, for another scale the two estimates are about *equal*, and on the third scale the NALS estimate is somewhat *lower*. Given the sampling errors associated with the small NALS subsample for this tri-county area, these results are consistent with the hypothesis that there are *not* substantial differences between the assessment methodologies of the two surveys. Once again, we find no evidence that the substantially higher statewide benchmark levels found for Oregon than found for other states or the nation as a whole are artifacts of methodological differences between the surveys used.

A final piece of evidence comes from an ongoing longitudinal study of a random sample of low-education adults in the Portland metropolitan area. The Longitudinal Study of Adult Learning (LSAL) is following a random sample of adults who were high school dropouts, age 18-44 and proficient (not necessarily native) speakers of English at the beginning of the study (Reder & Strawn, 2001). Adults' Document literacy skills were assessed with the Test of Adult Literacy Skills (TALS), another ETS-developed instrument which uses the same assessment methodology used in the NALS and the Oregon Literacy Survey. If we compare the assessed Document literacy proficiencies across the three studies for comparably defined populations (i.e., age 18-44, high school

dropout without GED, proficient English speaker), we find that the two Oregon studies find quite similar distributions of literacy proficiency, with much higher levels of literacy proficiency than found in other states or for the U.S. as a whole by the NALS. Although these comparisons focus on a subpopulation of adults, they provide further evidence that the results of the Oregon Literacy Survey are not anomalous – it is adult literacy proficiencies that are quite distinct in Oregon, not the methodology used to measure it.

Predicting Oregon’s Adult Literacy Benchmarks for 2000. Although the 1990 Census-based synthetic estimation models appear to do a reasonable job predicting NALS-measured county and state benchmark levels for much of the U.S., they substantially underestimate the state benchmark levels observed in the Oregon Literacy Survey of 1990. Our exploration of the differences between benchmark levels in Oregon and elsewhere in the U.S. indicate that adult literacy proficiencies are much higher in Oregon, even when differences in the demographic and educational characteristics of the populations are taken into account.

There is thus little reason to suppose that applying those same models to Oregon’s 2000 Census data would provide valid benchmark estimates either. For the sake of completeness, the statewide benchmark estimates for 2000 are displayed in Table 4. Corresponding estimates for individual Oregon counties are presented in Appendix C. Table 4 shows the model-generated benchmark estimates for both 1990 and 2000 to facilitate comparison. As can be seen, based on changes in Census between 1990 and 2000, the model predicts that Oregon’s benchmarks will increase slightly over the intervening decade. Given that these models do not accurately predict Oregon’s measured benchmark levels for 1990, we should have little confidence at this point in the increasing trend predicted between 1990 and 2000.

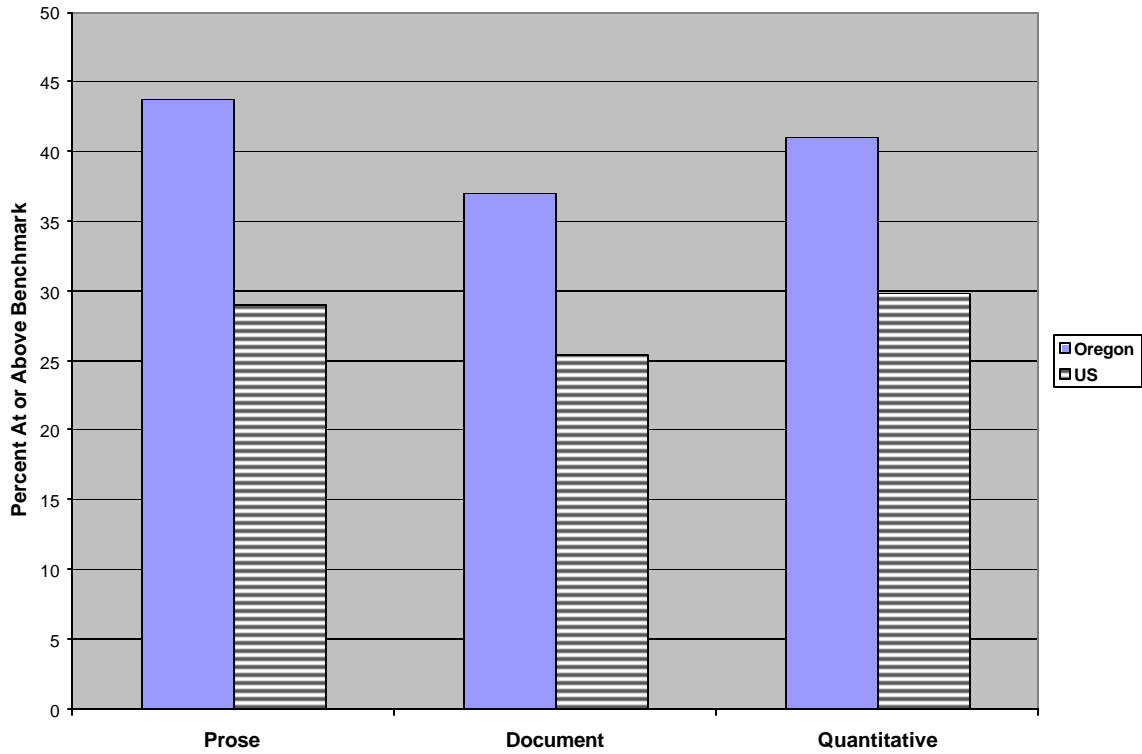


Figure 8. Literacy benchmark values for U.S.-born adults who are white, non-Hispanic and age 16-65. Source: Author calculations from the 1990 Oregon Literacy Survey (Oregon data) and the 1992 National Adult Literacy Survey (U.S. data).

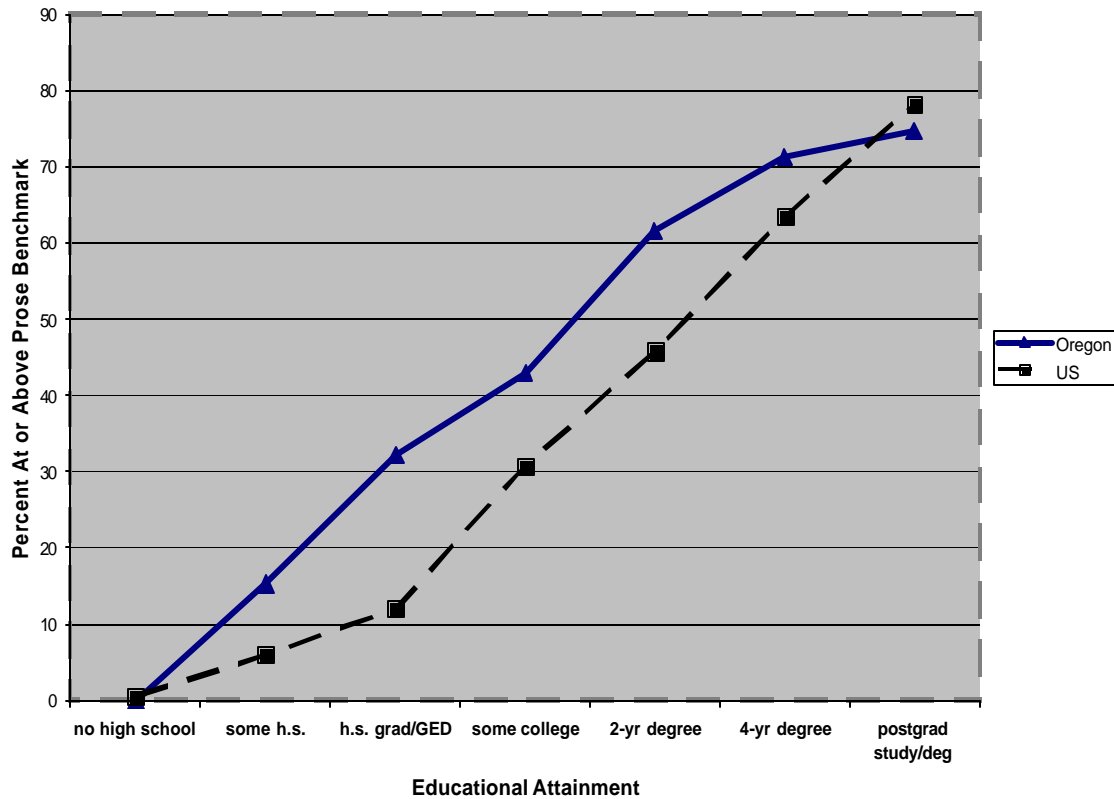


Figure 9. Percentage of Oregon (solid line) and US (dashed line) adult populations meeting or exceeding benchmark for Prose literacy proficiency as a function of educational attainment. Population comparison based only on adults age 16-65, US-born, white, and not Hispanic. Source: Author calculations from the 1990 Oregon Literacy Survey (Oregon data) and the 1992 National Adult Literacy Survey (U.S. data).

	PROSE	DOCUMENT	QUANTITATIVE
1990	30.2%	25.6%	30.1%
2000	31.0%	26.3%	30.5%

Table 4. Census-based predictions of adult literacy benchmarks for 1990 and 2000. Source: Author calculations from 1990 and US Census data and National Adult Literacy Survey. *These estimates are not believed to be valid.* See text for details.

Task 1 Summary. The synthetic estimation models developed in Task 1 that link the 1992 NALS and the 1990 U.S. appear to do reasonably well at job predicting county and state benchmark levels observed for much of the U.S. In general, the models do not predict the Oregon benchmark levels as accurately as earlier models predicted other benchmark levels, likely due to several methodological requirements for the earlier and current models: (1) the target Oregon benchmarks are modeled for a population age 16-

65 rather than for the entire adult population age 16 and above; (2) the target Oregon benchmark levels are for much higher proficiency levels than those predicted earlier; and (3) a smaller set of predictor variables are available for use jointly in the 1990 and 2000 models needed for Task 1.

Even though the models may perform satisfactorily for the U.S. as a whole, they substantially underestimate the state benchmark levels observed in the Oregon Literacy Survey of 1990. Our analysis of the differences between benchmark levels in Oregon and elsewhere in the U.S. indicates that adult literacy proficiencies are indeed much higher in Oregon, even when differences in demographic and educational characteristics of the populations being compared are taken into account.

These results have two major implications at this point in the project. First, Task 1 does not provide reliable estimates of Oregon's target adult literacy benchmarks. We may have better results in Task 2, which start with the adult literacy proficiencies observed during the 1990 Oregon Literacy Survey and then model the effects of a decade of population change on the benchmark levels.

The second implication is that Oregon may wish to reevaluate its adult literacy benchmark definitions and descriptions. It was noted above that Oregon's benchmark level of 325 and above (equivalent to Level 4 and above) is much higher than the national adult literacy benchmark level of 275 (Level 3 and above). This important definitional difference is not currently stated correctly in the Oregon Progress Board's benchmark publication and website and as a result fails to identify the much higher adult literacy levels observed in Oregon than in neighboring Washington State or in the U.S. in general. If it is decided to retain the 325 benchmark level, comparable data for Oregon and Washington are available in the "observed" bars in Figures 4, 5 and 6 for the Prose, Document and Quantitative scales, respectively. If it is decided to revise the benchmark to the 275 level, Table 5 displays comparable 275 benchmark levels for Oregon (Oregon Literacy Survey) and for eleven states and the United States as a whole (NALS restricted to individuals age 16-65). As expected, Oregon's benchmark levels are much higher than those of the U.S. and significantly higher than those of neighboring Washington State; although Washington State's 275 benchmark levels are similar to Oregon's.

	PROSE	DOC	QUANT
OREGON	78.0	76.1	80.0
UNITED STATES	59.1	55.8	58.7
CALIFORNIA	58.2	54.5	57.6
ILLINOIS	59.2	57.3	60.9
INDIANA	64.2	59.7	63.6
IOWA	68.7	66.1	72.3
LOUISIANA	50.1	44.8	48.5
NEW JERSEY	56.2	53.6	58.0
NEW YORK	51.7	49.6	53.0
OHIO	64.2	60.3	61.4
PENNSYLVANIA	62.3	59.9	64.0
TEXAS	52.9	50.3	51.8
WASHINGTON	74.6	71.1	73.2

Table 5. Estimates for 275 benchmarks for state and national populations, age 16-65. Source: Oregon: Oregon Literacy Survey; U.S. and other states: National Adult Literacy Survey.

Task 2

General approach. Although it may seem obvious that adult literacy scores in a state population are likely to change over time, it is not so simple to explication the reasons why scores would increase (or decrease over time). One way to approach the question of “why would adult literacy scores in Oregon increase between 1990 and 2000?” is to think of changes in two ways. One way that score can change is if the scores change for *individuals who reside in Oregon* in 1990 and 2000. If the average scores for residents increase, this will tend to raise the overall average for the state’s population. It does not mean, of course, that scores would need to change for every individual, but the individual gains would need to be larger that the individual losses.

For individual residents, there two major processes that occur over time. The first is the increasing education and concomitant literacy improvements for younger adults. In the early adult years, from the late teen years to approximately the late 20s, average literacy score increase because a relatively large proportion of younger adults are enrolled in college or other educational programs and are increasing their educational attainment.

The second process, and one that tends to counterbalance the first, is the diminishing literacy scores that occur as adults age. Literacy score tend to remain relative constant for individuals until about the mid-40s, after which they decline fairly steadily. We think that diminishing scores by age occur in the later adult years because a large number of adults do not spend much time reading, writing, or dealing with verbal or quantitative materials. This lack of practice is associated with literacy score reductions.

The second way in which adult literacy scores for Oregon could change is due to *individuals who move into and out of Oregon* during the 1990 to 2000 period. The entering and exiting individuals can influence the overall state's literacy scores if the average scores for those moving in (or moving out) differ from the average for the state's residents. During economic good times, when there are attractive employment opportunities in Oregon, there is usually in-migration of college-educated younger adults who increase the average adult literacy scores for the state. In an economically-weak period, out-migration can occur, which may tend to decrease the state's average literacy scores.

Taking the above two major processes into account, there are several factors that might increase the state's adult literacy scores over time:

- ? Increasing educational attainment among younger adults that improves their literacy scores,
- ? In-migration of better-educated adults who have literacy scores that are higher than the average for state residents, and
- ? Out-migration of less-educated adults who have literacy scores that are lower than the average for state residents.

Likewise, there are several factors that might decrease the state's adult literacy scores over time:

- ? Ageing of older adults that diminishes their literacy proficiencies,
- ? Out-migration of better-educated adults who have literacy scores that are higher than the average for state residents, and
- ? In-migration of less-educated adults who have literacy scores that are lower than the average for state residents.

Although the above factors emphasize the role of age and educational attainment, as discussed in later sections in this report, these are not the only factors that are associated with adult literacy proficiencies. For example, there are some relatively well-educated immigrants who move to Oregon and increase the state's average adult literacy scores. But, there are also a substantial number of immigrants who, largely because of their weak English-language skills, arrive in Oregon; they have the net effect of lowering the state's average adult literacy scores. Taking education and other factors into account, foreign-born individuals arriving in Oregon tend to lower the state's adult literacy scores. This is an example of one of many influences that affect the changes over time in the state's adult literacy benchmark values.

Population model. There are three ways to enter Oregon's adult population, aged 16 to 65 years: "ageing" from the teen years into the 16-65 age group, immigration, and in-migration from another state. And there are three ways to leave the same population age group: "ageing" from the adult years into the 66+ age group, emigration, and out-migration to another state. Each of these three entrance and exit routes raise the possibility of altering the state's average adult literacy scores. If the average scores for those entering the state's adult population, for instance, are higher than the average for

the current state's residents, this will raise the state's adult literacy scores. Likewise, if the average scores for those leaving the state are higher than the state's average, this will decrease the state's adult literacy scores. These effects can be taken into account in a population model for changes in the state's population over time.

There is another important process at work over time that also influences the state's adult literacy scores. Residents in the state do not have unchanging demographic characteristics. Residents become older over time. They may increase their educational attainment. And some adult residents die. We can take some of these processes into account in a population model. But, as explained later, we cannot model explicitly all possible changes that occur over time to the state's residents.

Data and methods. The population model is based on a statistical prediction of the average literacy scores – and the likelihood that literacy scores exceed a score of 325 – for the state's population. The procedure involves two major steps. The first step relies on the Oregon Literacy Survey, conducted in 1990 and based on household interviews with 1,993 adult Oregon residents. We analyze these survey data in order to develop a set of regression equations that predict the likelihood that an individual's score is greater than the benchmark value of 325. We include separate statistical analysis for each of the three Prose, Document and Quantitative components.

The second step uses 1990 and 2000 census data for Oregon, arrayed for a set of demographic characteristics that match the variables used in the regression analysis of the Oregon Literacy Survey. Using the 1990 and 2000 census, we predict the adult literacy benchmarks based on the regression equations.

There is a critical methodological caution to note about interpreting changes in adult literacy scores using this population model. The prediction equation is estimated from the 1990 Oregon Literacy Survey and there is no more recent survey data available to derive this prediction equation. *By applying the same prediction equation to 1990 and 2000 census data, we are able to take into account changes in Oregon's demographic characteristics. But we assume that the structure of literacy scores is the same within each of the demographic groups.* To give a specific example, the method predicts, in both 1990 and 2000, that someone who is native-born, white, aged 25 to 34 years, and a high school graduate will have a overall literacy score of 313 and 36.2 percent will have a score above 325. Of course, there are hundreds of possible subgroups and, even if there are small shifts within any group, they may counterbalance and any one group would not greatly affect the state's overall average. Nevertheless, it is possible that systematic shifts in adult literacy scores with many subgroups may have occurred in the 1990 to 2000 period and, if so, would have affected the prediction equation that we use for this work.

Regression analysis. The Oregon Literacy Survey was conducted in 1990 in conjunction with a larger national assessment of adult literacy. The Oregon survey involved in-person interviews with 1,993 adults. It collected demographic data from respondents and

included a detailed assessment of their literacy. In particular, the survey provides five scores dealing with prose, document, and quantitative ability – each of which include an average score based on an average of the five sub-scores. Finally, the survey includes an overall average literacy score for each respondent.

The observed variation in literacy scores for Oregon adults ranged from 170 to 389 for the overall literacy score in 1990. The average score for Oregon's was 309. Compared to the usual standard exceed a score of 325, 37 percent of Oregon adults had a score above 325.

We began the statistical analysis by examining variables that were collected in both the 1990 Oregon Literacy Survey and the U.S. decennial censuses for 1990 and 2000. We also based our examination on past research that suggested potential useful variables. For example, past research indicates convincingly that male/female gender does not predict adult literacy scores, taking other variables into account. We did not include gender in our analysis for this reason.

After reviewing data variables common to the Oregon Literacy Survey and the U.S. decennial census, we limited attention to the following variables:

- ? Age: defined in five ten-year age groups, including 16-24, 25-34, 35-44, 45-54, and 55-65;
- ? Race/Hispanic-Origin: defined as white or minority;
- ? Nativity: defined as native-born or foreign-born;
- ? Educational attainment: defined as 0-8 years of schooling, 9-11 years of schooling, high school graduate or GED, some college or two-year degree graduate, and college graduate or advanced degrees.

We considered using more detailed categories from some variables but found that the number of cases were too small for more categories. For example, it appears from national data that adult literacy scores vary for Asian Americans, blacks, Hispanics, and American Indians. There were relatively few respondents, however, for special racial/Hispanic groups in the 1990 Oregon Literacy Survey.

Each of the four variables is strongly related to adult literacy. Age is a strong predictor of adult literacy scores. Scores increase in the early adult years, when many individuals pursue advanced education, peak during ages 25 to 35, and decrease thereafter. Adult literacy scores are slightly lower for minority respondents than white respondents. Foreign-born respondents, primarily because their weaker English-language proficiency, have lower literacy scores than native-born respondents. Educational attainment has a consistent, strong association with adult literacy. Compared to respondents with 0-8 years of education, college graduate typically score more than 100 points higher on overall literacy scores.

Our preliminary analysis found that each of the above four variables had strong and statistically significant associations with adult literacy scores, based on the 1990 Oregon

Literacy Survey data. In addition, we found that there were three interactions that were important. Two variables are said to interact in affecting a prediction equation when the effects of one variable depends on the value of the other (Fox, 1997: 145-153). We decided to modify the regression equation to take these three interactions into account. The three interactions include:

- ? Age*Race: minority respondents do not display the same age profile for adult literacy scores as found for white respondents. All other factors being equal, minority respondents older than 25 years of age tend to score more poorly.
- ? Age*Nativity: foreign-born respondents do not show the same age profile for adult literacy scores as found for native-born respondents. All other factors being equal, foreign-born respondents tend to score more poorly in the 25-34 and 45-54 age groups and better in the 35-44 age groups than expected. These differences for foreign-born respondents are likely related to their age at the time of arrival in the United State and the duration of U.S. residence. The Oregon Literacy Survey, however, did not collect information on when foreign-born respondents arrived in the United States.
- ? Education*Race: minority respondents do not have the same educational profile for adult literacy scores as reported for white respondents. All other factors being equal, minority respondents with some college or a college degree have higher adult literacy scores than would be expected.

After examining several possible statistical models, the most accurate predictions came from a model that includes four main variables (age, race, nativity, and educational attainment) and three interactions (age*race, age*nativity, and education*race). For this analysis, four new variables were created for each respondent in the 1990 Oregon Literacy Survey: 0 if their literacy score (overall and for each of the three component scores) was less than 325, and 1 if their literacy score was 325 or higher. Because the variable to be predicted was binary (0-1), the variables were analyzed using logistic regression. In each case, the same explanatory variable was used as in the previous analysis. Based on logistic regression analysis, more than three-fourths (77 percent) of respondents could be correctly predicted in terms of whether their literacy scores were below or above 325. Appendix D presents the estimated logistic regression coefficients for the three prediction equations: one each for prose, document, and quantitative literacy components.

Census data. The second main step of the population model deals with preparing tabulations from the 1990 and 2000 U.S. Censuses of Population. The key feature of these tabulations is that the categories need to match exactly the definitions for the 1990 regression equations. In other words, the tabulations include the number of Oregon adults, aged 16 to 65 years, for five age groups, two race groups, two nativity groups, and five educational attainment groups.⁹ Moreover, the tabulations need to be constructed so that estimates take into account the three interaction terms for the regression equations.

⁹ The census microdata were available for age categories 18 to 64. An adjustment was made to take into account the number of Oregon residents aged 16 and 17 and age 65 in order to have comparable data with

We used census microdata available from the 1 percent Public-Use Microdata Sample (PUMS) for Oregon’s population in 1990 and 2000. These PUMS data allowed use to construct variables that were the same as those in the 1990 Oregon Literacy Survey. We then prepared tabulations that matched those required for use with the regression equations. This meant that we calculated the number of Oregon adult, in 1990 and 2000, who were in each of the 5 ages by 2 races by 2 nativity statuses by 5 educational attainment groups or, in other words, one of 100 possible groups.

The adult literacy estimates for Oregon, in 1990 and 2000, were derived by applying the regression equations to the known 1990 and 2000 Oregon census tabulations.

Results. Table 6 presents information on the literacy scores for Oregon’s adult population, aged 16 to 65 years, as observed in the 1990 Oregon Literacy Survey and as predicted from the population model using 1990 and 2000 U.S. census data. As seen in the second and third columns, the 1990 observed data and the 1990 predicted data are very similar, indicating that the predictive population model is fairly accurate and that there was probably very little change in literacy scores between 1990 and 1990.

	<i>Oregon Literacy Survey</i>	<i>Predicted Using 1990 and 2000 U.S. Census Data</i>			
	<i>1990</i>	<i>1990</i>	<i>2000</i>	<i>Difference</i>	<i>Percent Change</i>
Overall Literacy	37%	37%	36%	-1%	-3.71%
Literacy Components					
Prose	42%	42%	40%	-2%	-4.84%
Document	35%	35%	34%	-1%	-4.19%
Quantitative	39%	39%	38%	-1%	-3.34%

Table 6. Literacy Scores Exceeding 325 for Oregon Adults, Aged 16 to 65 Years: Observed in 1990 Oregon Literacy Survey; Predicted for 1990 and 2000 Censuses

the age range of 16-65 years in the Oregon Literacy Survey. The number of persons, in each separate tabulation for the age group 18-24 were multiplied times 9/7 to include ages 16 and 17 years. Furthermore, it was assumed that the additional 16 and 17 year-olds has some high school education. The results in this report are *not* sensitive to the assumption about their educational attainment: if all 16 and 17 year-olds had less education (only 0-8 years) or more education (high school degrees), it would change overall state adult literacy scores by less than 0.1. The number of persons, in each separate tabulation for the age group 55-64 were multiplied times 11/10 to include age 65 years. In addition, it was assumed that the additional 65 year-olds have the same educational attainment distribution as the 55-64 age group.

The predicted adult literacy benchmarks for 2000 indicate small and modest decreases in literacy during the 1990-2000 period. Benchmarks for the three components of literacy – prose, document, and quantitative – each display slight decreases.

Table 7 presents information for the percentage of Oregon’s adults with literacy scores above 325. There was a modest decrease the percent scoring higher than 325 for overall literacy, declining from 37 percent in 1990 to 36 percent in 2000, a percentage decrease of 3.7 percent. Similar modest decreases were observed for each of the three literacy components – prose, document, and quantitative.

In the midst of fairly substantial in-migration to Oregon, including a large number of college graduates, why does the population model show literacy score declines? We can disentangle the source of demographic change by using actual 1990 and 2000 census data and the original regression equations. We do this by performing a set of “what if” statistical predictions.

For the “what if” experiments, we start with the actual 1990 population of Oregon adults. First, we ask: what if Oregon’s population in 1990 had the age composition reported in 2000, but that all other factors remained unchanged? What would Oregon’s literacy score have been if only the age composition were different? Table 7 shows the results of this though experiment for each of the variables that we measure in the regression equations.

	<i>Predicted Mean Literacy Score</i>	<i>Predicted Literacy Score</i>
		325+
1990 Predicted Score	308.17	37.01%
2000 Predicted Score	305.98	35.63%
Overall Difference	-2.19	-1.37%
Difference Due to:		
Age	-0.74	-0.29%
Schooling	1.15	0.40%
Race	-1.69	0.02%
Nativity	-0.27	-0.04%
Age*Race	-0.42	-0.09%
Age*Nativity	-0.30	-0.38%
Schooling*Race	0.14	-0.83%
Other Interactions	-0.05	-0.18%

Table 7. Decomposition of Differences Between 1990 and 2000 Mean Literacy Scores and Percentages with Scores Exceeding 325 Based on Population Model Results

We know from Table 6 that the predicted literacy benchmarks decreased slightly by a percent or two. Figure 10 shows the observed shifts in Oregon’s age composition, including decreases in the 25-35 year-old age group and increases in the 45-54 age group. If only the 1990-2000 age composition changes were to have occurred, then the predicted percentage of adults exceeding the 325 benchmark would have diminished by .29 percentage points. What happened, in part, in the 1990s was that Oregon’s age composition became older and, by 2000, included a greater proportion of adults in age groups with lower literacy scores.

The effect of shifts in Oregon’s educational attainment, as noted in the differences due to schooling in Table 7, has the effect of increasing adult literacy benchmark scores. There were changes in the educational attainment groups for all groups, except for high school graduates who comprised about 31 percent of the state’s adult population in both 1990 and 2000. If the only change experienced in Oregon during 1990 to 2000 would have been the observed increases in educational attainment, then the state’s mean adult literacy

score would have increased by 1.15 points and the predicted percentage with scores exceeding 325 would have gained .4 percentage points.

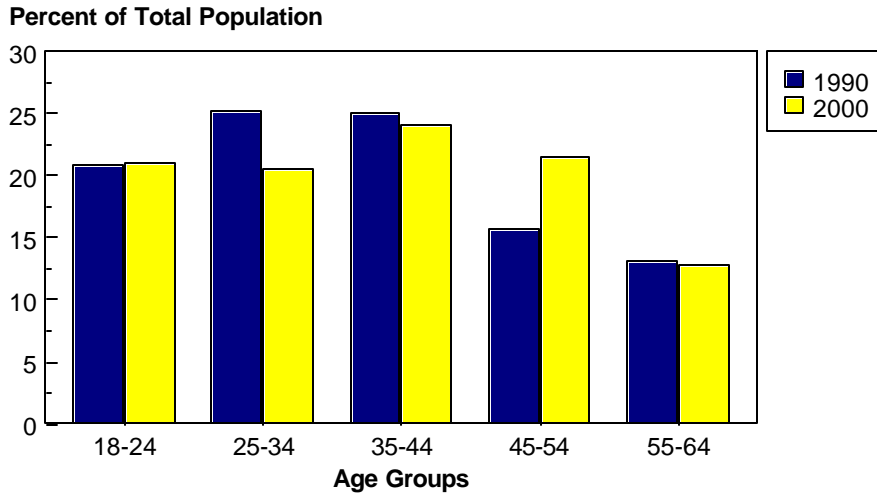


Figure 10. Percentage Distribution of Oregon's Population by Age Groups, 1990 and 2000

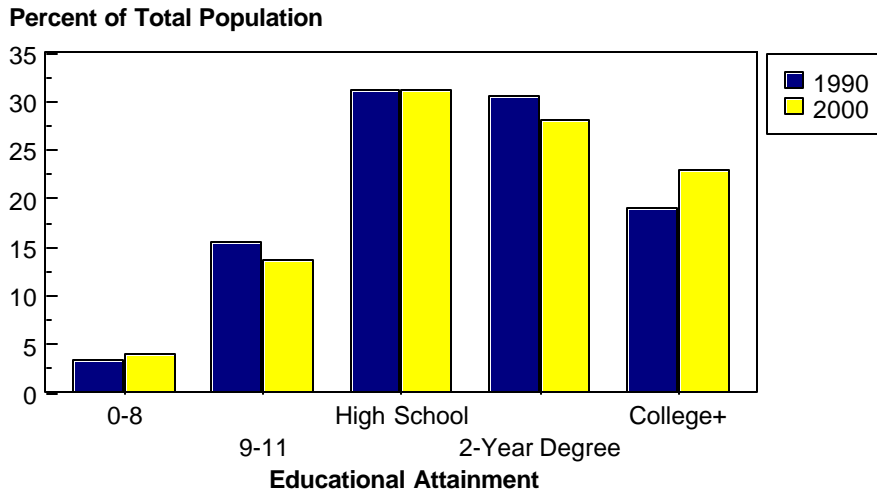


Figure 11. Percentage Distribution of Oregon's Population by Educational Attainment Groups, 1990 and 2000

Percent of Total Population

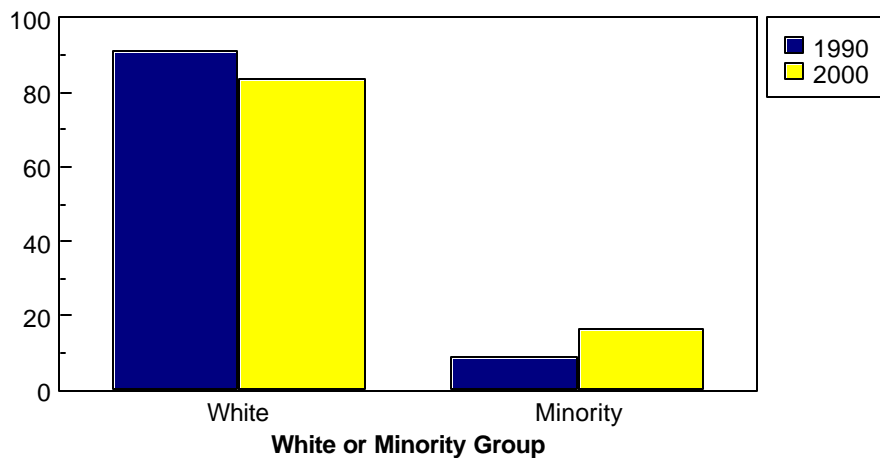


Figure 12. Percentage Distribution of Oregon's Population for White and Minority Groups, 1990 and 2000

Percent of Total Population

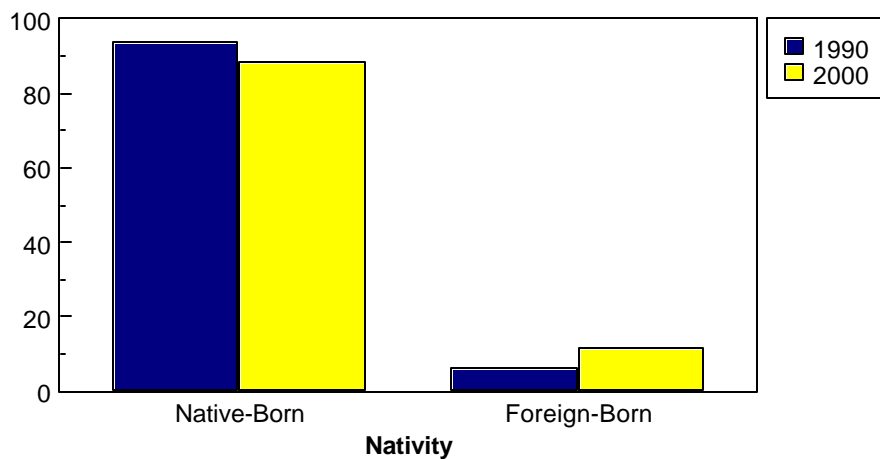


Figure 13. Percentage Distribution of Oregon's Population for Native-Born and Foreign-Born, 1990 and 2000

Percent of Total Population

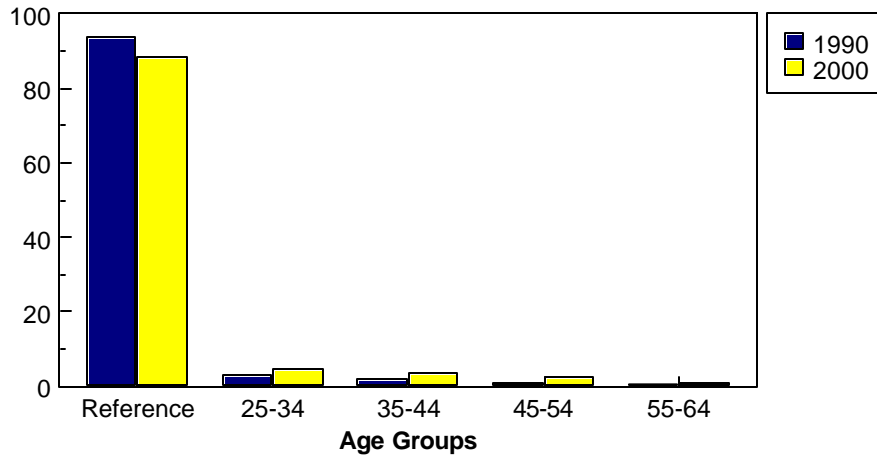


Figure 14. Percentage Distribution of Oregon's Population for Minority Population by Age, Compared to Remainder of the Population, 1990 and 2000

Percent of Total Population

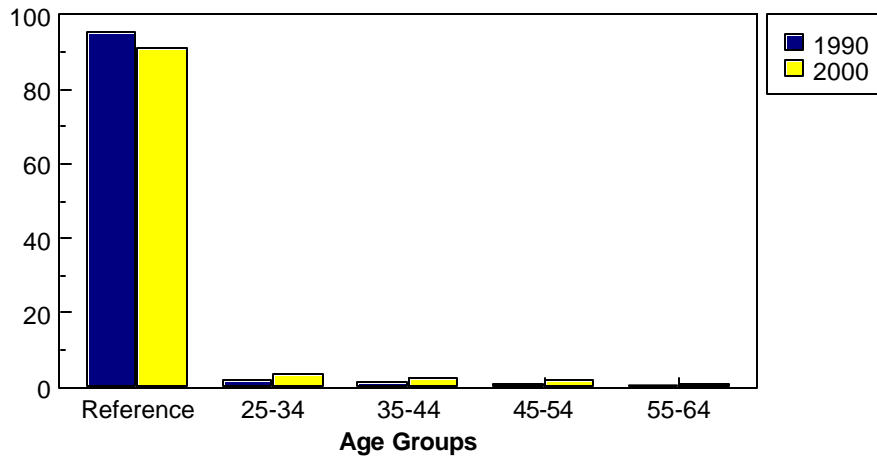


Figure 15. Percentage Distribution of Oregon's Population for Foreign-Born Population by Age, Compared to Remainder of the Population, 1990 and 2000

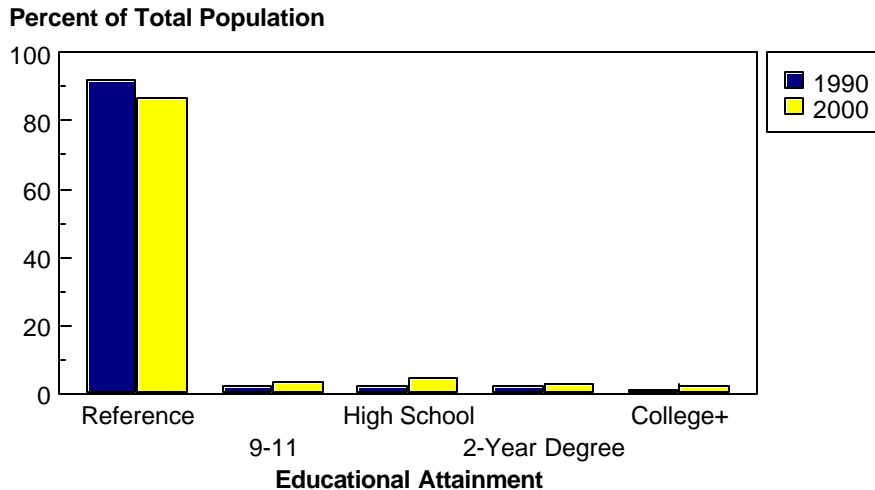


Figure 16. Percentage Distribution of Oregon's Population by Educational Attainment Groups for the Minority Population, Compared to the Remainder of the Population, 1990 and 2000

During the 1990-2000 period, as indicated in Figure 12, the state experienced substantial increases in its minority population. If this had been the only change during the decade, it would have influenced minor decreases in the predicted benchmark percentages. We hasten to note, as mentioned earlier, that the state's minority population is heterogeneous. It includes a large number of Asian Americans, including immigrants, who arrived in Oregon in the 1990s, are well-educated, and improved the state's literacy proficiency. But, primarily because we lack a survey with sufficiently large number of minority respondents, we are not able to include more detailed categories for racial/ethnic minorities.

The foreign-born generally have somewhat lower literacy scores, taking educational attainment and other factors into account, than native-born residents. As shown in Figure 15, there were increases in the proportion of the state's population who were foreign-born. If the only change during the 1990-2000 period were increases in the foreign-born population, this would have led to a very small decline in the percentages of the state's adult population with scores exceeding the benchmark level of 325.

There were changes in each of the three interactions included in the statistical model. Change in each of these three factors – the minority population by age, the foreign-born population by age, and the educational attainment of the minority population – are shown in Figures 14, 15, and 16 respectively. Increases in the minority population by age and in the foreign-born population by age, if no other changes had occurred in the state's demographic composition between 1990 and 2000, would have decreased the predicted benchmark value. Improvements in the educational attainment of the minority

population, on the other hand, would have increased the predicted percentages exceeding the benchmark values.

Overall, two factors in the 1990s contributed to improvements in the state's adult literacy benchmarks: increases in educational attainment and improvements in the educational attainment among the minority population. These gains, however, were counterbalanced by several trends: an ageing of the state's population, increases in the minority and foreign-born populations, and ageing of the minority and foreign-born populations. Combined together, these negative influences worsen slightly the demographic factors associated with predicted percentage with the predicted percentage above the literacy benchmarks. The changes were not profound, though, and the best evidence from the population model is that the state's adult literacy scores were similar, albeit slightly lower, in 2000 compared to 1990.

Task 2 summary. Task 2 developed a population model for estimating Oregon's adult literacy benchmarks in 1990 and 2000. The model is based on regression analysis of the Oregon Literacy Survey data from 1990 – a large household survey of 1,993 respondents who were asked a variety of questions and administered an assessment of their literacy proficiencies along three scales (prose, document, and quantitative).

Analysis of the 1990 survey revealed a regression model that offers a suitable prediction for an individual's literacy status. Models were developed to predict whether individuals exceeded the prose, document, and quantitative benchmark levels of 325. The models included five categories of age, white/minority status, foreign/native-born, and five categories of educational attainment. Further statistical analysis demonstrated the need to include three interaction terms because there were three combination of variables in which an individual's literacy scores depended upon the specific combination (the report describes these interaction terms).

Data from the 1990 and 2000 censuses of population were tabulated for Oregon adults for the same categories used in the regression analysis. The regression equations were used to predict Oregon's literacy levels in 1990 and 2000. Although the changes between 1990 and 2000 were modest, the population model indicates that the state's adult literacy benchmarks decreased slightly from 1990 to 2000.

More detailed statistical analysis shows that the slight declines in literacy scores were the result of several demographic changes: the ageing of the state's population, increases in the minority and foreign-born populations, and the ageing of the state's minority and foreign-born populations. On the other hand, the improved educational attainment of adults offset some of these influences and, had no other factors affected adult literacy, would have improved the state's adult literacy benchmarks.

Summary of Tasks 1 and 2

Looking at both Task 1 and Task 2 together, we have tenuous results. Task 1 produced estimates for Oregon's adult literacy benchmarks for 1990 and 2000 based on models linked to the U.S. Census for those two years and to the National Adult Literacy Survey of 1992. These cross-sectional models show *a very slight increase* in Oregon's predicted literacy benchmarks from 1990 to 2000. Although these models do a good job predicting adult literacy levels for a variety of counties and states across the U.S., they do not perform satisfactorily at predicting the literacy proficiencies measured directly in the 1990 Oregon Literacy Survey.

In Task 2, on the other hand, the modeling began with the state's literacy proficiencies directly surveyed in 1990 and used a population model to project likely changes in literacy levels as a byproduct of changing population characteristics in Oregon between 1990 and 2000. Overall, the population models predict *very small decreases* in adult literacy benchmark levels between 1990 and 2000.

At the onset of the project, we stated that moderate and consistent results from the two approaches would offer a strong indication that Oregon's adult literacy scores had changed in the 1990-2000 period – either up or down. There is, however, no evidence that the levels changed substantially during the 1990s. This is indeed the juncture at which we have arrived. *In order to substantiate changes in adult literacy levels since 1990, for the overall state population and for population subgroups, an assessment based on a literacy survey is needed.*

One point of agreement between the two models is that neither predicts much change in Oregon's adult literacy benchmarks between 1990 and 2000. Although both models associate increasing levels of educational attainment in the adult population with improved literacy, the effects of other demographic changes in the adult population over the decade appear to offset those gains, resulting in a very modest net change in benchmark levels (a slight increase in one model, a slight decrease in the other). Although it is not surprising that the population model closely predicts Oregon's 1990 observed benchmark levels (since it started with those data), it is of considerable interest that the Task 1 models do not come close to predicting those models based on similar literacy data collected from the U.S. as a whole. As our report illustrates, this discrepancy appears to be the direct result of there being *substantially* higher levels of adult literacy in Oregon than in the rest of the country. Although this is potentially very good news for Oregon, and news we recommend that the Progress Board publicize accurately in its reports and web pages, we are still left here reluctant to project updated adult literacy benchmark estimates until more recent literacy data are collected. We therefore turn to Task 3 which will explore some cost-effective ways of conducting literacy assessments as part of the ongoing Oregon Population Survey.

Task 3

Purpose of Task 3. Task 3 of the project explores technical issues, design considerations and costs for a number of alternative ways of conducting statewide adult literacy assessments to provide periodic assessments of the state benchmark levels in literacy. The first section below considers both add-on approaches, in which an adult literacy component would be added to interviews already being conducted as part of an existing statewide survey, and free-standing approaches in which an entirely independent survey is conducted to assess the adult literacy benchmarks. In comparing the add-on and free-standing survey options, it is clear that the add-on approaches, if feasible, are more cost-effective. By capitalizing on the planning, sampling and background interviewing that are already done as part of an existing survey, an add-on approach to assessing adult literacy proficiency is the more cost-effective method for obtaining periodic updates to the adult literacy benchmark levels. The final section below considers the most promising of these options in more detail, an add-on component to the biannual Oregon Population Survey.

Options for conducting adult literacy assessments. In considering the various options for assessing adult literacy levels in Oregon, it is helpful to note the key factors that should be weighed. First is cost: how much would the assessment cost? Second is the scientific validity and reliability: would the assessment provide scientifically accurate and useful data? And third and final is the timeliness of the assessment: would it provide information every few years or only once a decade?

The requirements for adult literacy assessments include the following: it should provide a fairly accurate estimate of adult literacy, for all Oregon residents as well as for selected demographic groups (for age and sex groups, for example). There are three ways, we think, for conducting adult literacy assessments: (a) population and statistical models, (b) in-person surveys, and (c) add-on telephone surveys. We review each of these three approaches next.

Population and statistical models. This approach is illustrated by the analysis described in the first two tasks in this report. The approach is relatively inexpensive because it relies on existing census and survey data. Because the population model requires detailed cross-tabulated census data, this approach can be conducted only once a decade when census data are available. The regression model used for Task 1 here illustrates the usefulness of a statistical model approach. This approach also relies on census data, which are available only every ten years. Overall, population and statistical models offer a reasonably inexpensive approach for the assessment of adult literacy. Their major limitations are, however, that they require census data. Moreover, as seen earlier, these models are somewhat insensitive to minor changes in adult literacy and may not provide consistent estimates for changes in adult literacy.

In-person surveys. The use of in-person surveys has been the standard approach for assessment of adult literacy. This approach was used for the Oregon Literacy Survey in 1990, the National Adult Literacy Survey in 1991, and the 2003 National Assessment of

Adult Literacy that is currently underway. These surveys provide data with high validity and accuracy that can be cross-tabulated for selected demographic groups. Furthermore, once the initial costs of the survey are made, it is relatively inexpensive to include additional questions on the survey. This means that an additional advantage for this approach is that there is the possibility of a broad range of content in the survey questionnaire. The approach is relatively expensive (about \$725,000 or more for a survey for the state of Oregon) and, primarily because of high costs, is likely to be conducted only once every ten years.

Add-on telephone survey. One potential attractive approach for the assessment of adult literacy is to consider including a set of supplemental questions on an existing telephone survey. This approach makes two key assumptions: that a suitable telephone survey is conducted and that it is feasible to include a short set of questions, by telephone, to include for the assessment of adult literacy. Such an approach would provide timely data – if suitable surveys are conducted on a regular basis – and would be reasonably inexpensive. In Oregon, the Oregon Population Survey offers a potentially ideal survey because it is conducted statewide on a regular basis. The main challenge for this approach is that phone survey assessment of adult literacy is not well tested. An add-on telephone survey, at least for the initial implementation, would require an independent validation in order to document the accuracy of the survey results.

Exploring the feasibility of an add-on to the Oregon Population Survey. The previous section identified an add-on module to the Oregon Population Survey (OPS) as the most promising and cost-effective approach to measuring the adult literacy benchmark levels. This final section of the report examines this option in more detail. We first describe how this add-on would work. We next describe some existing instrumentation that might be used for assessing adult literacy in an OPS add-on. A critical step for planning such an add-on, a validation study, is then discussed. Some technical design issues that need to be considered in planning a validation study are next considered. Finally, a rough timeline and budget for developing and conducting the add-on component are discussed.

How the Oregon Population Survey add-on would work. The Oregon Population Study (OPS), a telephone survey of the Oregon population, is conducted every two years. It is next scheduled to be conducted starting in fall of 2004. The add-on literacy assessment module would be administered to a randomly selected subsample of OPS respondents. Selected respondents would be asked, immediately after completing the base OPS interview, to participate in an additional interview that would take about 30 minutes. They would be offered an incentive fee for this additional participation. If respondents agreed to participate, they would be interviewed immediately, if possible, or at another scheduled time when they would be called back. By sequencing the recruiting and interviewing in this fashion, the add-on component would not affect the participation rate or responses given in the main OPS. Furthermore, the additional costs of the add-on component would be limited to the costs of (1) the extra time of interviewers for administering the add-on questions; (2) extra time spent recontacting respondents by phone who had agreed to participate in the add-on but could not do so immediately; and

(3) the incentive fees for participation in the add-on. We estimate these costs would likely total an average of about \$25 - \$30 per add-on participant.

Instrumentation for assessing literacy by telephone. A key issue to be addressed in designing a literacy add-on to a telephone survey such as OPS concerns the kinds of questions or instrumentation that can be administered by telephone that would provide valid measures of the adult literacy benchmark. To some, it may seem implausible to begin with that literacy proficiencies – which obviously are based on the use of written materials – could be assessed by telephone at all. The approach suggested here, which is supported by a small but rigorous body of research, is based on administering questions by telephone for which correct answers serve as *indicators* or *proxies* of the types of literacy skills and knowledge assessed directly by the Oregon Literacy Survey (and related surveys such as the NALS that use the same instrumentation). Although the use of such indirect measures would certainly not be appropriate for all assessment purposes, for the purpose of estimating the literacy benchmark rates for the state’s adult population, it may be feasible to construct valid measures that can be orally administered by telephone.

Recent research suggests that such techniques may indeed be feasible. West, Stanovich and Mitchell (1993) demonstrated that engagement in literacy practices is highly correlated with measures of both literacy-related *skills* and literacy-related *knowledge*. They showed that simple checklist measures of literacy-related knowledge -- such as an author recognition test in which adults identify which names on a checklist are well-known authors – are highly correlated with scores on a range of standardized reading assessments. The theoretical base underlying such relationships was well articulated by Stanovich and associates for children and adults (Stanovich, 1993; Stanovich & Cunningham, 1993; West et al, 1993) and then further developed and applied by Sticht and colleagues for adults (Hofstetter, Sticht & Hofstetter, 1999; Sticht, 2000; Sticht, Hofstetter & Hofstetter, 1996).

Although the original “checklist” studies conducted by Stanovich and colleagues were conducted face-to-face with written materials, Sticht and associates adapted these instruments and methods for oral administration and used them in random digit dialing telephone surveys (Hofstetter et al, 1999; Sticht et al, 1996). The studies of the Sticht group established that these orally administered measures (1) are appropriate for telephone administration; (2) produce similar results when administered orally by telephone or in written form by mail; and (3) exhibit the same basic patterns of variation as standardized adult literacy assessments when results are disaggregated by background characteristics such as educational attainment, age, gender, race and ethnicity.

Four instruments were developed and used by Sticht and colleagues in their telephone surveys: the (1) author recognition test, adapted from the one used by West et al (1993); (2) a similar magazine recognition test; (3) a vocabulary test; and (4) a “cultural literacy” test. These four instruments take a total time of about 30 minutes on the average to administer by telephone, so they might well be appropriate for use in an add-on in the

Oregon Literacy Survey. The instruments themselves and the procedures for administering them by telephone and scoring them are published in Sticht et al (1996).

Need for validation study. Although previous research suggests that these telephone-administered instruments could be usefully incorporated into a range of adult literacy assessment designs (Sticht, 2000), their suitability for any particular adult literacy assessment purpose such as assessing Oregon's literacy benchmark levels should be carefully examined. A concurrent validation study is essential to evaluate their validity for this purpose and to calibrate their scores against the Oregon literacy benchmarks. In such a validation study, a large group of adults would be given *both* the telephone-administered instruments and the Oregon Literacy Survey assessment instrument or an equivalent instrument such as the commercially available Test Of Adult Literacy Skills (TALS). By comparing individuals' performance across both instruments, it could be determined how accurately the telephone-administered instruments could measure the Oregon literacy benchmarks. This information could be used to (1) decide whether to proceed with incorporating the add-on into the Oregon Population Survey; (2) optimize the subsampling design for the add-on to obtain the most efficient benchmark estimates; and (3) estimate the statewide benchmark levels from the add-on assessment data subsequently collected by the Oregon Population Study. If the validation study indicated in step (1) that the add-on would not be a valid or cost-effective measure of the benchmark levels, then steps (2) and (3), of course, not be carried out.

The validation study itself probably is best designed as a small random digit dialing survey that collects basic demographic information from individuals and administers the add-on assessment module. Respondents in the validation study should be paid to participate in the telephone assessment and then come to one of several local assessment centers (e.g., a nearby community college) for a group-administered paper-and-pencil literacy assessment (e.g., with the TALS). We would recommend paying participants in the validation study a stipend of about \$40 for both the telephone and face-to-face activities, to be paid upon completion of the group testing session. Many surveying costs can be minimized in the validation study because it does not require a carefully drawn and recruited random sample, but rather more of a purposive sample drawn from different regions of the state, different educational backgrounds and from a mix of gender, race and ethnic groups. A sample of about 400 - 500 would likely be adequate for the purposes of the validation study.

Timeline and rough budget for planning and conducting the OPS add-on assessment.

- ? February – April, 2004: plan and conduct validation study
- ? May 2004: analyze results of validation study
- ? June 2004: decide, based upon analysis of validation study, whether to proceed with adding-on assessment component to the Oregon Population Survey
- ? June 2004: design (sub)sampling plan for add-on
- ? July 2004 and beyond: fold add-on assessment component into design, training and implementation of Oregon Population Survey

Here is a rough budget for planning and implementing the add-on: (Note: This is provided for informational purposes only, not as a bid or proposal to conduct this work):

Validation Study

Design and management of validation study: \$5,000

Sampling \$2,500

Implementation of validation study: \$60 per respondent x 400 = \$24,000

(includes labor for contacting and interviewing respondents and respondent incentive fees)

Analysis of validation study and design recommendations for OPS add-on: \$5,000

Subtotal for validation study: \$36,500

Implementation of add-on to OPS (assuming sample size of 1,000 for add-on)

Additional interviewer training: \$1,500

Additional CAPI programming: \$1,500

Respondent incentives: \$15 x 1,000 = \$15,000

Additional interviewer labor: average \$10 per respondent x 1,000 = \$10,000

(includes additional interviewing labor per respondent and labor for calling back respondents unable to complete add-on immediately after base OPS interview)

Data analysis and reporting of state benchmark levels: \$5,000

Additional management costs for add-on: \$3,000

Subtotal for add-on (n=1,000): \$36,000

References

- Fox, J. (1997) *Applied Regression Analysis, Linear Models, and Related Methods*. Thousand Oaks, CA: Sage Publications.
- Hofstetter, C. R., Sticht, T. G., & Hofstetter, C. H. (1999). Knowledge, literacy, and power. *Communication Research*, 26, 58-90.
- Kaestle, C. F., Campbell, A., Finn, J. D., Johnson, S. T., & Mikulecky, L. J. (2001). *Adult literacy and education in America: Four studies based on the National Adult Literacy Survey*. Washington, DC: National Center for Education Statistics.
- Kirsch, I. S., Jungeblut, A., Jenkins, L., & Kolstad, A. (1993). *Adult literacy in America: A first look at the results of the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- National Educational Goals Panel (1997). *The National Education Goals report: Building a nation of learners*. Washington, DC: U.S. Government Printing Office.
- Oregon Progress Board (2003). *Is Oregon making progress? The 2003 Benchmark Performance Report*. Report to the Oregon Legislative Assembly, March 2003. Salem, OR: Author.
- Oregon Progress Board (1991). *The Oregon Literacy Survey: Measuring adults' functional skills*. Salem, OR: Author.
- Reder, S. (1998). Literacy selection and literacy development: Structural equation models of the reciprocal effects of education and literacy. In M. C. Smith (Ed.). *Literacy for the 21st century: Research, policy, practices, and the National Adult Literacy Survey*. (pp. 139-157). Westport, CN: Greenwood Publishing (Praeger).
- Reder, S. (1996). *Synthetic estimates of literacy proficiency for small Census areas*. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.
- Reder, S., & Edmonston, B. (2000). *Demographic change and literacy development in a decade*. Washington DC: National Center for Education Statistics.
- Reder, S. & Strawn, C. (2001). The K-12 school experiences of high school dropouts. *Focus on Basics*, 4(D), 12-13.
- Stanovich, K.E. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. In H. Reese (Ed.), *Advances in child development and behavior*, Vol. 24. New York: Academic Press.
- Stanovich, K.E., & Cunningham, A.E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, 85, 211-229.

Sticht, T. G. (2000). *Using telephone and mail surveys as a supplement or alternative to door-to-door surveys in the assessment of adult literacy*. Washington DC: National Center for Education Statistics.

Sticht, T. G., Hofstetter, C. R., & Hofstetter, C. H. (1996). Assessing adult literacy by telephone. *Journal of Literacy Research*, 28, 525-559.

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, 28, 35-50.

Appendix A

Regression equation predicting % at or above PROSE benchmark

	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
	B	Std. Error					
(Constant)	.029	.069		-.417	.677		
ED18_12	.146	.086	.096	1.691	.092	.233	4.296
ED18_HS	.114	.077	.105	1.479	.140	.149	6.706
ED18_CO	.277	.079	.176	3.498	.001	.295	3.388
ED18_AA	.366	.118	.106	3.103	.002	.638	1.567
ED18_BA	.754	.087	.405	8.677	.000	.344	2.907
ED18_GR	1.032	.088	.528	11.782	.000	.374	2.674
BLACK	-.181	.021	-.292	-8.813	.000	.686	1.458
HISPANIC	-.117	.044	-.164	-2.678	.008	.201	4.966
ENGNOTW	.216	.127	.127	1.698	.090	.135	7.397
ENGWELL	-.111	.094	-.055	-1.172	.242	.342	2.924
RECENT_I	-.273	.158	-.088	-1.724	.086	.289	3.455
REGION1	-.061	.011	-.226	-5.398	.000	.428	2.339
REGION2	-.013	.011	-.051	-1.175	.241	.395	2.534
REGION3	-.033	.011	-.136	-3.033	.003	.375	2.670

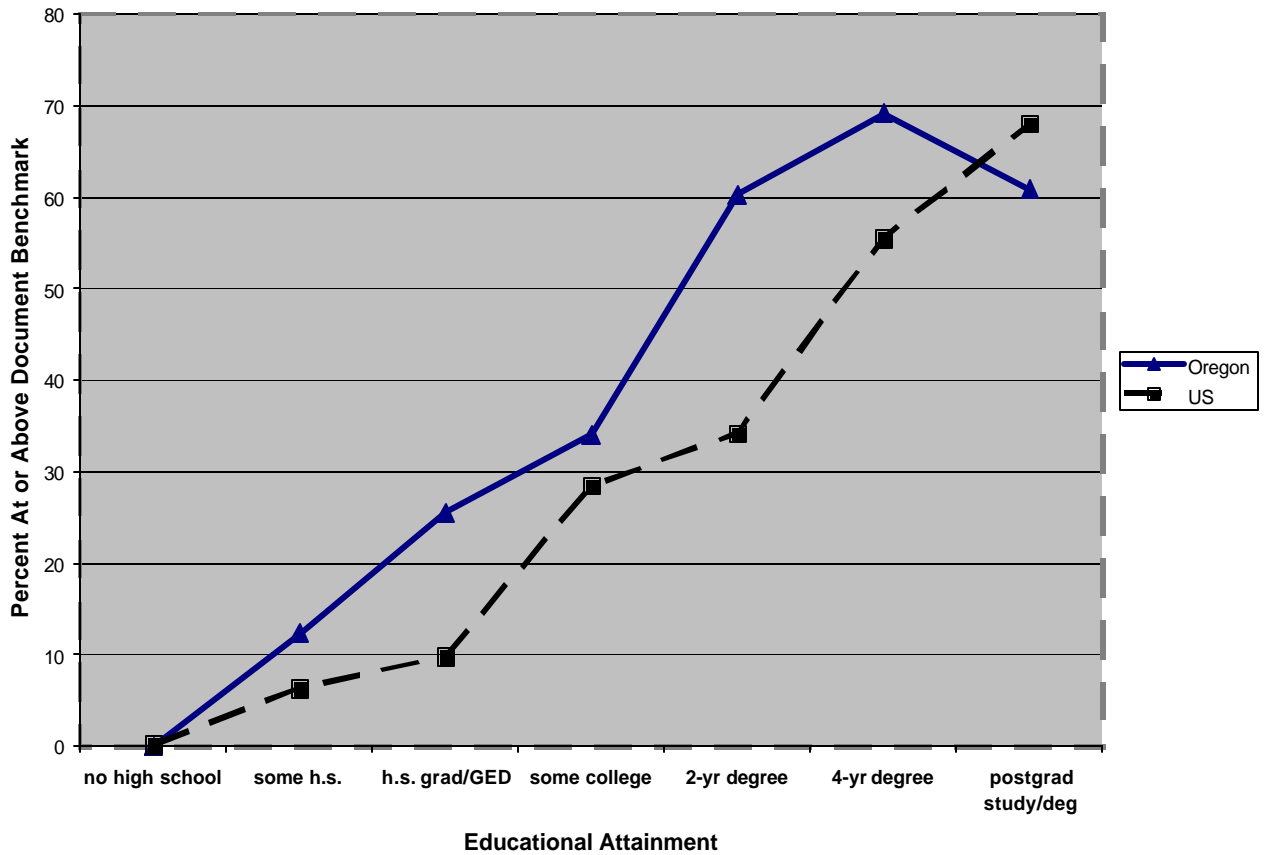
Regression equation predicting % at or above DOCUMENT benchmark

	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
	B	Std. Error					
(Constant)	-.036	.066		-.556	.578		
ED18_12	.133	.083	.099	1.601	.110	.212	4.711
ED18_HS	.094	.073	.094	1.283	.200	.152	6.578
ED18_CO	.300	.075	.213	3.981	.000	.285	3.510
ED18_AA	.322	.114	.099	2.820	.005	.671	1.490
ED18_BA	.560	.083	.338	6.765	.000	.328	3.053
ED18_GR	.928	.082	.521	11.297	.000	.385	2.595
BLACK	-.211	.020	-.378	-10.612	.000	.647	1.546
HISPANIC	-.120	.038	-.203	-3.191	.002	.202	4.958
ENGNOTW	.269	.113	.185	2.385	.018	.136	7.375
ENGWELL	-.206	.086	-.120	-2.408	.017	.330	3.030
RECENT_I	-.135	.142	-.051	-.952	.342	.290	3.450
REGION1	-.029	.011	-.119	-2.739	.006	.436	2.292
REGION2	.007	.010	.032	.723	.470	.417	2.396
REGION3	-.021	.010	-.099	-2.081	.038	.361	2.774

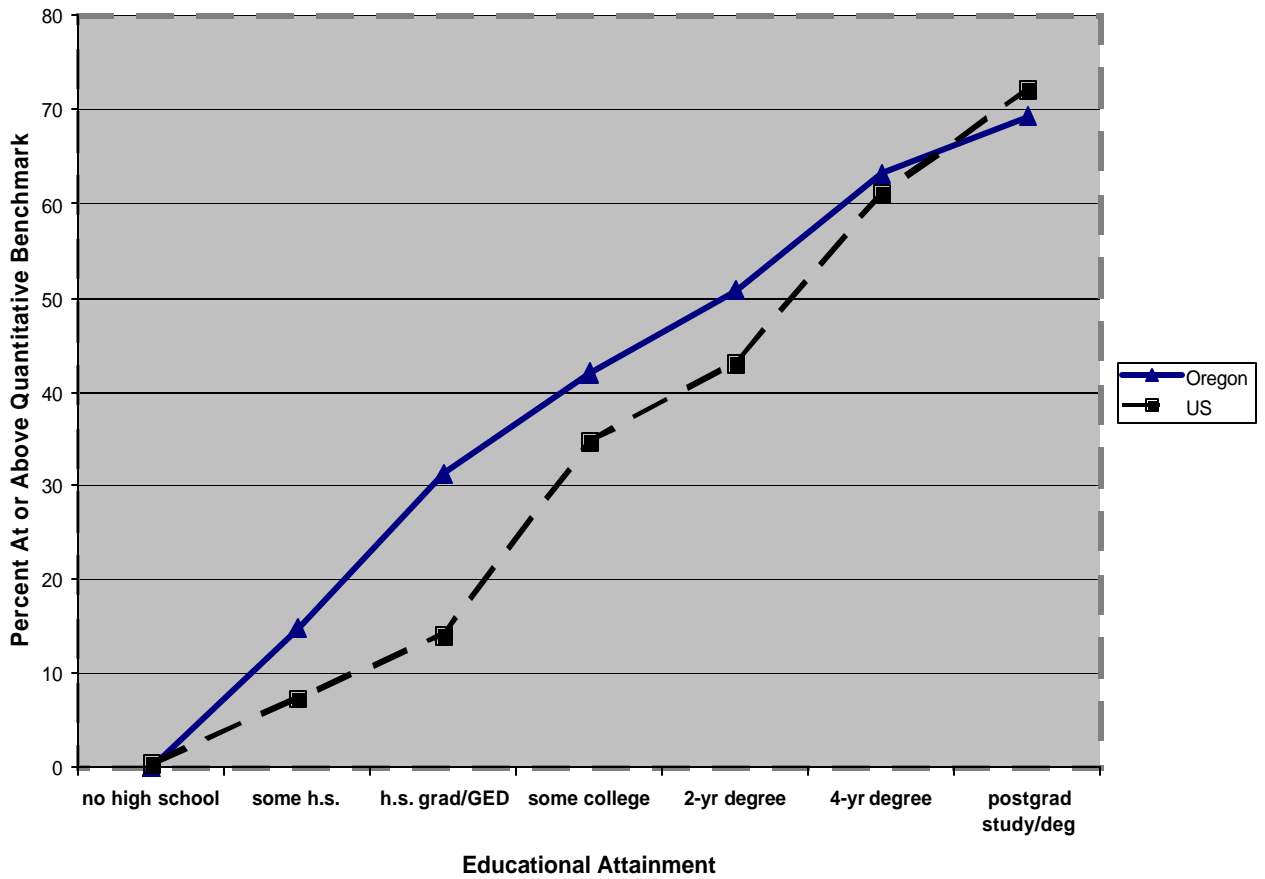
Regression equation predicting % at or above QUANTITATIVE benchmark

	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
	B	Std. Error					
(Constant)	-.000	.069		-.006	.995		
ED18_12	.043	.088	.028	.490	.624	.232	4.311
ED18_HS	.112	.078	.099	1.430	.154	.159	6.291
ED18_CO	.289	.080	.186	3.634	.000	.290	3.453
ED18_AA	.467	.124	.127	3.768	.000	.664	1.505
ED18_BA	.592	.086	.320	6.892	.000	.352	2.837
ED18_GR	.980	.087	.507	11.223	.000	.373	2.683
BLACK	-.215	.021	-.336	-10.052	.000	.682	1.467
HISPANIC	-.082	.042	-.120	-1.963	.050	.203	4.919
ENGNOTW	.062	.123	.038	.504	.615	.135	7.423
ENGWELL	-.179	.091	-.093	-1.980	.048	.344	2.907
RECENT_I	-.146	.150	-.049	-.970	.333	.303	3.304
REGION1	-.037	.011	-.137	-3.302	.001	.443	2.260
REGION2	.009	.011	.035	.820	.413	.420	2.382
REGION3	-.016	.011	-.065	-1.459	.145	.380	2.633

Appendix B



Percentage of Oregon (solid line) and US (dashed line) adult populations meeting or exceeding benchmark for Document literacy proficiency as a function of educational attainment. Population comparison based only on adults age 16-65, US-born, white, and not Hispanic. Source: Author calculations from the 1990 Oregon Literacy Survey (Oregon data) and the 1992 National Adult Literacy Survey (U.S. data).



Percentage of Oregon (solid line) and US (dashed line) adult populations meeting or exceeding benchmark for Quantitative literacy proficiency as a function of educational attainment. Population comparison based only on adults age 16-65, US-born, white, and not Hispanic. Source: Author calculations from the 1990 Oregon Literacy Survey (Oregon data) and the 1992 National Adult Literacy Survey (U.S. data).

Appendix C *

COUNTY	PROSE	DOCUMENT	QUANTITATIVE
Baker	26%	23%	26%
Benton	44%	37%	43%
Clackamas	34%	29%	34%
Clatsop	27%	24%	28%
Columbia	25%	21%	26%
Coos	26%	22%	26%
Crook County	22%	19%	23%
Curry County	27%	23%	27%
Deschutes	32%	27%	32%
Douglas	24%	20%	25%
Gilliam	25%	22%	27%
Grant	26%	22%	26%
Harney	23%	19%	23%
Hood River	26%	22%	25%
Jackson	29%	25%	29%
Jefferson	21%	18%	21%
Josephine	25%	22%	26%
Klamath	25%	21%	25%
Lake	25%	21%	25%
Lane	32%	28%	33%
Lincoln	29%	25%	29%
Linn	24%	21%	25%
Malheur	18%	15%	18%
Marion	25%	22%	25%
Morrow	19%	16%	19%
Multnomah	34%	28%	32%
Polk	31%	27%	32%
Sherman	28%	23%	28%
Tillamook	25%	22%	25%
Umatilla	24%	21%	24%
Union	30%	25%	30%
Wallowa	30%	25%	30%
Wasco	25%	21%	26%
Washington	36%	30%	35%
Wheeler	25%	21%	25%
Yamhill	26%	22%	26%

* As explained in the text, these estimates are not believed to be valid. They are appended here only for the completeness of the report.

Appendix D

Table D-1. Logistic Regression Coefficients for Prediction of Overall, Prose, Document, and Quantitative Literacy Scores That Exceed 325, Oregon Literacy Survey, 1990

		<i>Overall</i>	<i>Prose</i>	<i>Document</i>	<i>Quantitative</i>
Constant		-7.435	-6.663	-6.215	-6.905
Age					
	<25	0.000	0.000	0.000	0.000
	25-34	0.075	0.173	-0.302	0.170
	35-44	-0.032	0.090	-0.474	0.345
	45-54	-1.090	-0.703	-1.496	-0.381
	55-64	-1.281	-0.845	-2.022	-0.364
Race					
	White	0.000	0.000	0.000	0.000
	Other	0.073	0.602	-0.010	-0.359
Nativity					
	Native-born	0.000	0.000	0.000	0.000
	Foreign-born	-0.156	-0.692	0.202	-0.606
Education					
	0-8	0.000	0.000	0.000	0.000
	9-11	4.758	4.935	4.071	4.826
	12	6.794	6.189	5.908	6.311
	2-Year Degree	7.550	6.727	6.672	6.716
	College+	9.350	8.397	7.917	7.811
Age*Race					
	Reference	0.000	0.000	0.000	0.000
	25-34, Other	0.029	-0.496	-0.811	0.379
	35-44, Other	-0.138	-0.833	-1.273	-0.182
	45-54, Other	-0.135	-1.200	-0.783	0.838
	55-64, Other	-3.895	-5.084	-4.566	1.024
Age*Nativity					
	Reference	0.000	0.000	0.000	0.000
	25-34, Foreign-Born	-0.795	0.223	-1.150	0.066
	35-44, Foreign-Born	-0.497	0.318	0.196	-0.049
	45-54, Foreign-Born	-7.123	-1.066	-6.206	-0.969
	55-64, Foreign-Born	-0.569	1.129	0.423	0.205
Education*Race					
	Reference	0.000	0.000	0.000	0.000
	9-11, Other	-5.397	-5.624	-0.617	-1.327
	12, Other	-3.272	-1.535	-3.030	-1.684
	2-Year Degree, Other	-1.264	-2.495	-0.354	-0.876
	College+, Other	-1.865	-2.111	-0.316	-0.729
Summary					
	Nagelkerke R squared	0.425	0.340	0.362	0.264
	% Predicted Correct	77.100	72.200	74.800	70.600