

Wyoming's Industry and Occupational Projections Methodology 2010 to 2020

Tony Glover, Analyst

*"The only function of economic forecasting is to make astrology look respectable."**John Kenneth Galbraith*

The importance of this article lies in the fact that the data generated by the processes described are used by policy makers, educators, employers, job seekers, and others to anticipate future; training needs, educational curriculums, human resource and staffing needs, and career choices. Given these circumstances Research & Planning (R&P) is obligated to produce the most accurate and transparent projections possible. This article is divided into four parts and each part is a step in R&P's creation of Wyoming's Industry and Occupational Projections. Along the way differences between the current methodology and previously published projections are discussed, as well as, the reasons for revising our approach to projections. Part I describes the collection of base period data and the creation of detailed industry and sub-state region projections. Part II introduces a method to estimate job replacement needs due to permanent exits as a result of retirements, deaths, and exiting Wyoming's labor force. Part III discusses combining data from Parts 1 & 2 with data from the Occupational Employment Statistics program to create occupational projections. Part IV lists future avenues of research related to forecasting that R&P will explore over the next few years. For example, R&P is currently collecting data on a representative sample of newly hired employees in Wyoming which provides insight into the types of jobs employers are currently filling.

Part I - Baseline Industry and Sub-State Region Projections

*"The major task of time-series analysis is to describe the nature of the past variation of a variable in order that its future values can be predicted and acted upon accordingly."**(Kachigan, 1986)*

There are qualitative as well as quantitative approaches to forecasting. In most cases the forecasting models involve a mixture of both approaches. There are many sophisticated models such as autoregressive moving average (ARMA), autoregressive integrated moving average

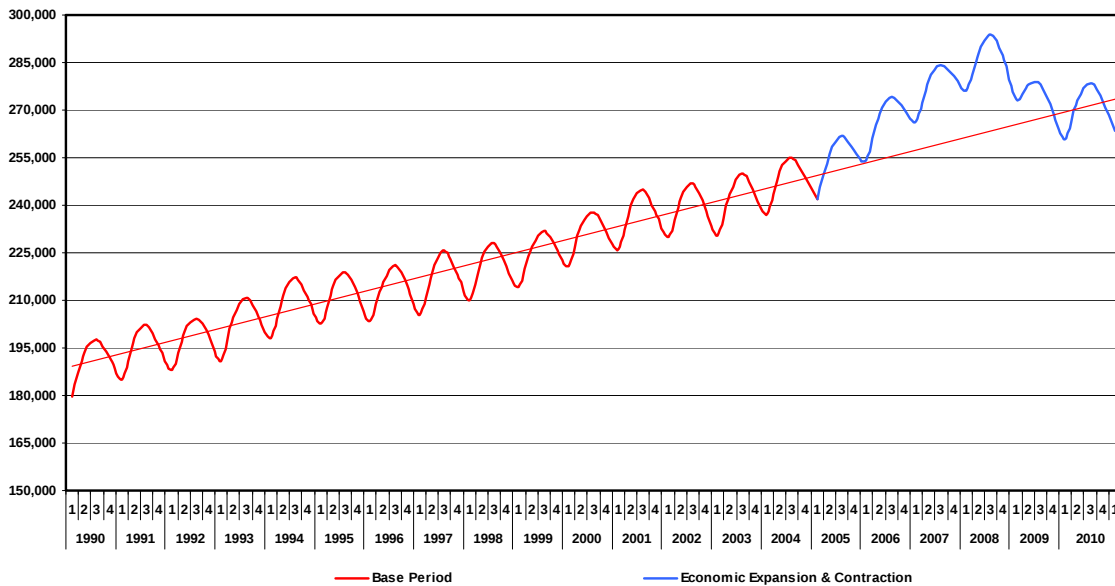
(ARIMA, e.g. Box-Jenkins), extrapolation, and linear prediction to name a few. Regression modeling techniques allow the introduction of predictor variables, such as, oil and gas prices, new home construction permits, etc. R&P has used many of these approaches in the past and often combined the approaches by applying one model to industry A and another to industry B. Predictive models make an assumption that the factors influencing the predictor variables are known. J. S. Armstrong, the founder of the International Institute of Forecasters and the International Journal of Forecasting, stated “*Sophisticated extrapolation techniques have had a negligible payoff for accuracy in forecasting. ... Meanwhile, simple methods and the combination of forecasts are recommended*” (Armstrong, 1984). In 2011, Armstrong reiterated this statement “*Do not use complex methods; they do not improve accuracy and the added complexity can cause forecasters to overlook errors and to apply methods improperly. We do not recommend complex econometric methods.*” (Armstrong, 2011). The content in Part I demonstrates that simple linear regression techniques yielded forecasts similar to, and slightly more accurate than, those created using more sophisticated modeling techniques and requiring more resources (analyst and time).

As the world moves toward a global economy, markets are no longer closed systems and the economic principle of ‘all else being equal’ no longer applies. Natural gas prices over the past decade were not only driven by increasing market demand but were greatly influenced by pipeline construction. Natural gas pipelines made far flung resources easily available in global markets. Internet based book stores, such as Amazon.com, and online streaming video stores, such as NetFlix.com have virtually eliminated local book and video stores. Recent investments in energy efficiency and healthcare will change the direction of two of Wyoming’s largest industries and have unforeseen consequences for indirectly impacted industries.

Figure 1.1 shows the average quarterly employment in Wyoming from 1990q1 to 2011q1. A quick glance reveals that from 1990 to 2004 Wyoming had a consistent continuous period of growth. The rapid growth from 2005 to 2008 resulted in dramatic increase in forecasted employment both in the short-term and long-term industry and occupation projections. The straight line in Figure 1.1 is a simple linear regression based on the period 1990q1 to 2004q4 and

it suggests that the recent recession in Wyoming may have been a course correction and the dramatic growth from 2005 to 2008 was an anomaly.

Figure 1.1: Average quarterly employment in Wyoming from 1990q1 to 2011q1



A portion of R&P's funding comes from the Workforce information Grant (WIG) which requires R&P to produce long-term and short-term industry and occupational projections during alternating years. Over the past decade R&P has done this by utilizing the Projections Management Partnerships (PMP) projections software which incorporates quantitative aspects of forecasting by presenting the analyst with several of the models mentioned above and allowing the introduction of other predictive time series. It also addresses qualitative aspects by allowing the analysts to decide which forecasting model best fits the data based on their knowledge of the economy, as well as, allowing the analyst to change the series as they see fit. Simply put the projections created are not reliable because on any given day R&P's M.A. level engineer will disagree with R&P's M.A. level economist over the direction a specific industry in a designated sub-state area will go. The projections are also difficult to defend as the software rolls the detailed industry and sub-state region projections, based on different models and assumptions, into aggregate data for presentation purposes. This aggregation strategy without documentation makes it difficult to determine how the pieces impact the whole.

The projection methodology discussed in this article uses a systematic approach that creates the same result given the same base period data and does not use the PMP software. The current model uses a combination of linear regression models (see formula 1.1) based on the last 120 months @ 80.0% weight and the last 36 months @ 20.0% weight of employment data at the 3 digit - sub-state region level. Several variations of the current model were run using different

Formula 1.1: Simple linear regression and standard error of the estimate.

$$\text{Slope (b)} = (N \sum XY - (\sum X)(\sum Y)) / (N \sum X^2 - (\sum X)^2)$$

$$\text{Intercept (a)} = (\sum Y - b(\sum X)) / N$$

$$\text{Regression Equation (y)} = a + b x$$

$$\text{Standard Error of the Estimate} = (\sum (y - y')^2 / n - 2)^{1/2}$$

projection horizons (i.e. last 120 months and last 24 months, last 120 months and last 12 months) and different weighting schemes (i.e. 120 months @ 90.0% and last 12 months @ 10.0%) but the model 120 months @ 80.0% & 36 months @ 20.0% (referred to as 80/20 blend for the remainder of the paper) produced the best predictive results. The evaluation of the blend combinations which included 74,760 data points (89 - 3-digit industries, 7 - sub-state regions, and 120 - months of data) for each run was done for every month from 2000 to 2010. The results of the previous long-term projections by major industry compared to the results of the 80/20 blend are in Table 1.1.

Table 1.1 shows the base year, projected year, and net change in employment for three of R&P's previously published long-term projections. Table 1.1 also includes the base year, projected year, and net change in employment from the 80/20 blend model using the data that was available when the previously published projections were created. Lastly, the far right columns apply the projections forecast to estimate what employment would be in 2010 and the percentage error between the estimated 2010 employment and the actual 2010 employment. The values for the actual employment in 2010 are available as the base year actual employment in the bottom panel of Table 1.1 (2010 to 2020 projections series).

A review of Table 1.1 shows that while both models overestimated the employment in 2010 the 80/20 blend was closer to the actual employment in 2010 in two of the three (2004 to 2014 & 2006 to 2016) projections presented. The third set of models, for 2008 to 2018, overestimated the employment in both cases and 80/20 blend was slightly less accurate than the projections published by R&P using the PMP software.

In comparing the two approaches we need to point out a couple of differences between the determination of the employment in the base year and the timeliness (when the projections were created relative to data availability) of the two projection models.

The first difference is the calculation of average employment for the base year. Previously published projections represent the average monthly employment during the base year and the 80/20 blend uses the average employment of November from the year prior to the base year and May of the base year. This was done in the 80/20 blend to normalize the base year employment to the Occupational Employment Statistics programs benchmarked employment, for reasons discussed in Part III of this document, the occupational projections methodology, which follows.

Secondly, as the 80/20 blend is systematic and was run on data available as of May in the base year with no adjustments to the model based on data subsequent to May. The development of the previously published projections involved several analysts and a longer period of time (a considerable investment in time and money) as each industry and region graph was reviewed. In the case of the 2008 to 2018 projections Quarterly Census of Employment and Wages (QCEW) data were already available for at least the first two quarters of 2009 when the projections were completed. A review of Figure 1.1 allows you to ascertain the impact of this knowledge on the analysts qualitative input to the previously published models (in their own words they wrenched the projections down to be as conservative as possible).

The final 2010 to 2020; total all industries, by major industry, and by sub-state region projections are presented in Figures 1.2, 1.3, and 1.4 at the end of this part. Figure 1.2 to 1.4 shows the actual base employment from 2000 to 2010 and the projected employment from 2011 to 2020. Figure 1.2 also shows the 95% confidence interval of the projected data based on the variation between

the actual employment data from 2000 to 2010 and the back casted regression line. The Standard Errors of the Estimates (see Formula 1.1) were created at the 3-digit NAICS and sub-state region level and aggregated to the total as presented in Figure 1.2. The error estimates are not shown in Figures 1.3 and 1.4 but are calculated and applied to the occupational projections in Part III.

In R&P's PY 2010 Workforce Information Grant Performance Report we introduced the Employment and Training Administration (ETA) to our new projection methodology. This article is in response to a request for our methodology. It also sparked a lively conversation with ETA about whether R&P had considered other advantages of the PMP software for which the ETA has invested \$4 million dollars in development. Three major issues for R&P are; First, the systematic process used in creating the projections discussed in this section are reliable, in that the same results are produced every time given the same set of data. Secondly, the approach is at least as accurate as prior processes which required a greater investment of analyst time. Lastly, the data used is already prepared as part of the QCEW program and projections are run in minutes rather than months which make it timelier and self-correcting as projections can be generated every quarter when the QCEW is updated.

References for Part I

Kachigan (1986). Statistical Analysis – An Interdisciplinary Introduction to Univariate & Multivariate Methods. Radius Press, New York, NY 1986.

Armstrong (2011). Demand Forecasting: Evidence-Based Methods, Oxford Handbook in Managerial Economics, pending publication.

Armstrong (1984). Forecasting by Extrapolation: Conclusions from 25 Years of Research. Interfaces, 14 (Nov.-Dec), 52-56, 1984.

Table 1.1: Industry Projection Model Performance for Three Projection Horizons

		Projection Models						Projected Estimate of 2010 Employment vs. Actual 2010 Employment					
		Published Projections			80% & 20% Blend *			Published Projections			80% & 20% Blend		
Projection Series	NAICS Industry	Base Year Actual Employment	Projected	Net	Base Year Actual Employment	Projected	Net	Proj2010	Projected-Actual	Percent Error	Proj2010	Projected-Actual	Percent Error
2004to2014	00 Total	248,032	311,919	63,887	245,317	289,643	44,326	286,364	15,878	5.9%	271,913	1,426	0.5%
2004to2014	11-21 Natural Resources & Mining	22,529	34,217	11,688	21,483	28,196	6,713	29,542	2,847	10.7%	25,511	-1,184	-4.4%
2004to2014	42, 48-49 Wholesale Trade, Transportation, & Utilities	17,460	23,859	6,399	16,700	20,351	3,652	21,299	1,576	8.0%	18,890	-833	-4.2%
2004to2014	23 Construction	19,177	27,904	8,727	19,948	26,435	6,487	24,413	1,267	5.5%	23,840	693	3.0%
2004to2014	31-33 Manufacturing	9,536	11,326	1,790	9,595	10,202	607	10,610	1,839	21.0%	9,959	1,187	13.5%
2004to2014	44-45 Retail Trade	29,951	34,468	4,517	29,787	32,498	2,711	32,661	2,939	9.9%	31,414	1,691	5.7%
2004to2014	51 Information	4,938	5,609	671	4,278	5,355	1,078	5,341	1,452	37.3%	4,924	1,036	26.6%
2004to2014	52-53 Financial Activities	10,562	13,473	2,911	10,354	12,699	2,345	12,309	1,466	13.5%	11,761	919	8.5%
2004to2014	54-56 Professional & Business Services	15,699	21,131	5,432	15,580	19,122	3,542	18,958	2,121	12.6%	17,705	867	5.2%
2004to2014	61-62 Education & Health Care	53,861	65,296	11,435	52,856	65,813	12,958	60,722	-512	-0.8%	60,630	-604	-1.0%
2004to2014	71-72 Leisure & Hospitality	32,601	38,655	6,054	29,422	29,706	284	36,233	5,968	19.7%	29,592	-673	-2.2%
2004to2014	81 Other Services	7,970	9,447	1,477	7,385	8,294	909	8,856	1,030	13.2%	7,930	104	1.3%
2004to2014	90 Government (Includes Postal Service)	23,748	26,534	2,786	27,932	30,974	3,043	25,420	-6,115	-19.4%	29,757	-1,778	-5.6%
2006to2016	00 Total - All Industries	266,860	324,452	57,592	261,744	325,002	63,259	289,897	19,410	7.2%	287,047	16,560	6.1%
2006to2016	11-21 Natural Resources & Mining	28,737	35,794	7,057	27,073	40,721	13,649	31,560	4,865	18.2%	32,532	5,837	21.9%
2006to2016	22, 48-49 Transportation, Warehousing, & Utilities	11,174	14,988	3,814	10,555	13,787	3,232	12,700	1,555	13.9%	11,848	703	6.3%
2006to2016	23 Construction	24,011	32,433	8,422	23,511	35,151	11,640	27,380	4,233	18.3%	28,167	5,020	21.7%
2006to2016	31-33 Manufacturing	10,126	11,042	916	10,009	10,710	701	10,492	1,721	19.6%	10,289	1,518	17.3%
2006to2016	42 Wholesale Trade	8,292	10,366	2,074	8,046	10,952	2,906	9,122	544	6.3%	9,208	631	7.4%
2006to2016	44-45 Retail Trade	30,999	34,545	3,546	30,651	34,702	4,051	32,417	2,695	9.1%	32,271	2,549	8.6%
2006to2016	51 Information	4,711	5,021	310	4,029	4,372	343	4,835	947	24.3%	4,166	277	7.1%
2006to2016	52-53 Financial Activities	11,192	13,788	2,596	10,918	14,076	3,158	12,230	1,388	12.8%	12,181	1,338	12.3%
2006to2016	54-56 Professional & Business Services	17,449	24,837	7,388	16,829	21,044	4,215	20,404	3,567	21.2%	18,515	1,678	10.0%
2006to2016	61 Educational Services	25,079	28,926	3,847	26,767	30,886	4,120	26,618	-3,338	-11.1%	28,414	-1,542	-5.1%
2006to2016	62 Health Care & Social Assistance	28,971	36,800	7,829	27,324	34,914	7,591	32,103	825	2.6%	30,360	-918	-2.9%
2006to2016	71-72 Leisure & Hospitality	33,960	40,368	6,408	30,268	32,745	2,477	36,523	6,258	20.7%	31,259	993	3.3%
2006to2016	81 Other Services	8,556	10,288	1,732	7,710	9,683	1,974	9,249	1,422	18.2%	8,499	672	8.6%
2006to2016	90 Government (Including Postal Service)	23,603	25,256	1,653	28,057	31,262	3,205	24,264	-7,271	-23.1%	29,339	-2,196	-7.0%
2008to2018	00 Total, All Industries	286,317	312,050	25,733	283,744	347,183	63,439	291,464	20,977	7.8%	296,431	25,945	9.6%
2008to2018	11-21 Natural Resources & Mining	31,696	31,523	-173	30,522	45,280	14,758	31,661	4,967	18.6%	33,474	6,779	25.4%
2008to2018	22, 48-49 Transportation, Warehousing, & Utilities	12,335	13,097	762	11,919	15,996	4,077	12,487	1,342	12.0%	12,734	1,589	14.3%
2008to2018	23 Construction	28,184	30,408	2,224	28,632	37,159	8,527	28,629	5,482	23.7%	30,337	7,191	31.1%
2008to2018	31-33 Manufacturing	10,056	9,638	-418	10,186	10,090	-96	9,972	1,201	13.7%	10,166	1,395	15.9%
2008to2018	42 Wholesale Trade	9,110	9,300	190	8,963	11,769	2,806	9,148	571	6.7%	9,524	946	11.0%
2008to2018	44-45 Retail Trade	32,312	33,751	1,439	32,156	35,517	3,361	32,600	2,877	9.7%	32,828	3,106	10.4%
2008to2018	51 Information	4,699	4,782	83	4,010	3,910	-100	4,716	827	21.3%	3,990	101	2.6%
2008to2018	52-53 Financial Activities	11,727	12,761	1,034	11,507	14,696	3,190	11,934	1,091	10.1%	12,144	1,302	12.0%
2008to2018	54-56 Professional & Business Services	18,871	21,851	2,980	18,426	21,291	2,865	19,467	2,630	15.6%	18,999	2,162	12.8%
2008to2018	61 Educational Services	26,313	30,344	4,031	28,048	33,846	5,798	27,119	-2,837	-9.5%	29,208	-748	-2.5%
2008to2018	62 Health Care & Social Assistance	31,372	38,926	7,554	29,313	39,876	10,563	32,883	1,605	5.1%	31,426	148	0.5%
2008to2018	71-72 Arts, Entertainment, Leisure & Hospitality	36,061	39,090	3,029	32,299	35,057	2,759	36,667	6,401	21.2%	32,850	2,585	8.5%
2008to2018	81 Other Services	8,805	9,560	755	8,194	10,095	1,901	8,956	1,130	14.4%	8,574	748	9.6%
2008to2018	90 Government (Including Postal Services)	24,777	27,019	2,242	29,571	32,604	3,033	25,225	-6,310	-20.0%	30,178	-1,358	-4.3%
2010to2020	00				270,487	307,108	36,622						
2010to2020	11-21				26,695	36,017	9,323						
2010to2020	22, 48-49				11,145	14,497	3,352						
2010to2020	23				23,147	24,092	946						
2010to2020	31-33				8,772	8,433	-339						
2010to2020	42				8,578	10,280	1,702						
2010to2020	44-45				29,723	29,463	-260						
2010to2020	51				3,889	3,597	-292						
2010to2020	52-53				10,843	12,340	1,498						
2010to2020	54-56				16,838	17,775	938						
2010to2020	61				29,956	36,152	6,196						
2010to2020	62				31,278	39,593	8,316						
2010to2020	71-72				30,266	31,598	1,333						
2010to2020	81				7,827	9,069	1,242						
2010to2020	90				31,535	34,205	2,670						

* 80% & 20% Blend combines 80% of the past ten years linear regression with 20% of the last three years linear regression at the three digit NAICS sub-state region which are aggregated to major industry in the table above. Several combinations were tested on historic data for example, 90% last ten 10% last two, but the model presented above performed the best over the range of economic business cycles.

Figure 1.2. Wyoming Monthly Employment for Base Period 2001 to 2010 & Projected Period 2011 to 2020 with 95% Confidence Interval.
Actual employment increase 2000 to 2010 - 14.3%
Projected employment increase 2011 to 2020 - 11.8

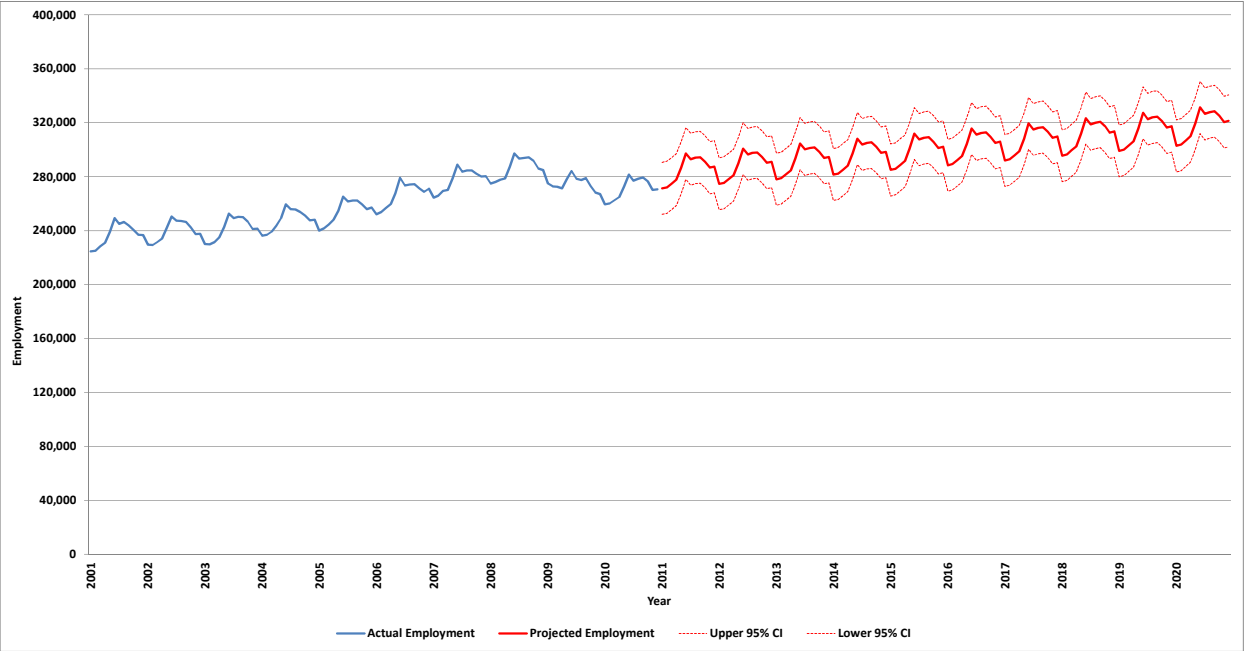


Figure 1.3. Wyoming Monthly Employment for Base Period 2001 to 2010 & Projected Period 2011 to 2020 by Major Industry.

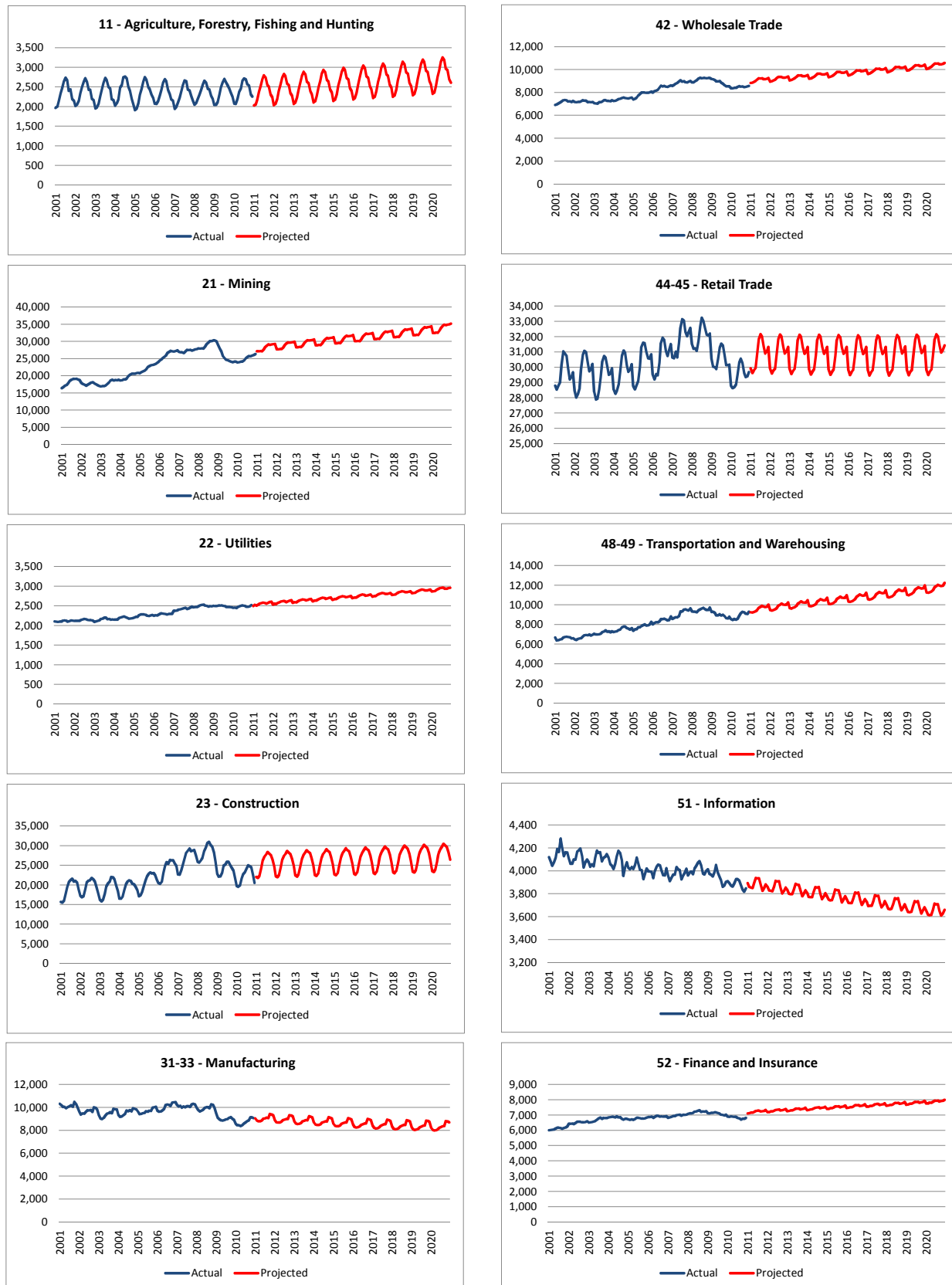


Figure 1.3. Wyoming Monthly Employment for Base Period 2001 to 2010 & Projected Period 2011 to 2020 by Major Industry.

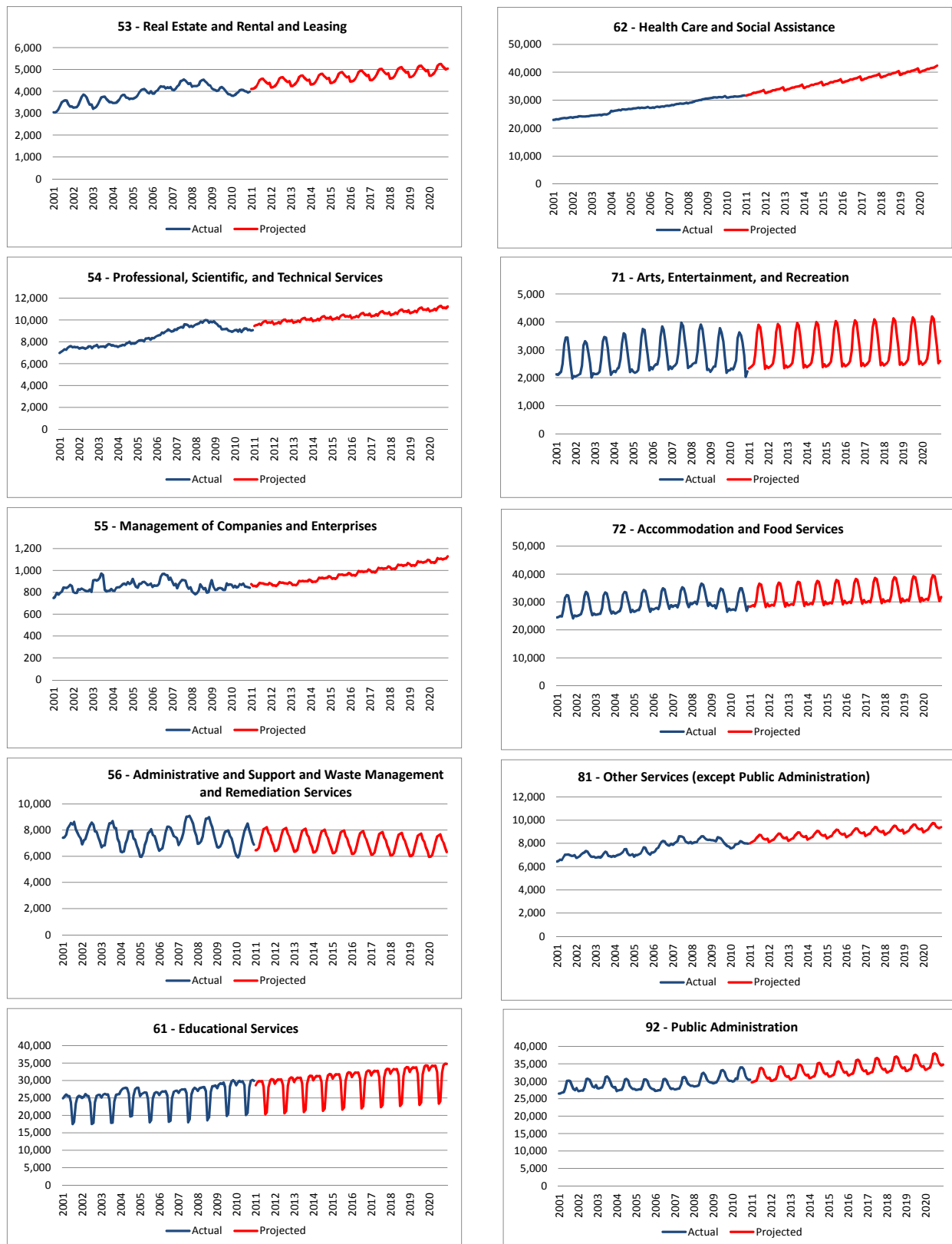


Figure 1.4. Wyoming Monthly Employment for Base Period 2001 to 2010 & Projected Period 2011 to 2020 by Sub-State Region.

