Linear Regression Analysis: Medicaid's Effect on U.S. BMI

Regression analysis is an important statistical method for the analysis of medical data. It enables the identification and characterization of relationships among multiple factors. It also enables the identification of prognostically relevant risk factors and the calculation of risk scores for individual prognostication.

BMI in 1990 and BMI in 2010
There is a positive correlation between BMI in 1990 versus BMI in 2010. BMI progressively increases from the 1990 to the 2010 time period.
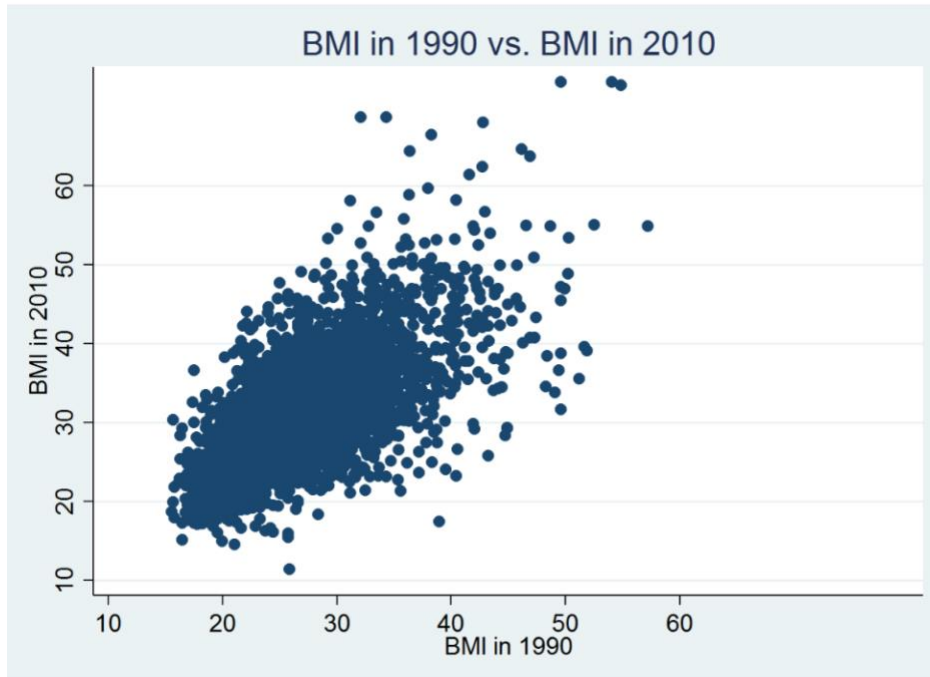
Baseline - BMI, Medicaid, and Health Insurance in 1990
This regression indicates that an individual on Medicaid in 1990 would have a BMI 1.698 points higher than someone who was not on Medicaid. Medicaid's effect on BMI is statistically significant. Private health insurance is not statistically significant.

```
. regress BMI90 YMed90 YHI90
```

Source	SS	df	MS		Number of obs	=	9,797
					F(2, 9794)	=	42.51
Model	2074.77426	2	1037.38713		Prob > F	=	0.0000
Residual	239030.8	9,794	24.4058404		R-squared	=	0.0086
					Adj R-squared	=	0.0084
Total	241105.575	9,796	24.6126556		Root MSE	=	4.9402

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.697922	.2060361	8.24	0.000	1.294049	2.101795
YHI90	.0179061	.1237397	0.14	0.885	-.2246493	.2604615
_cons	25.29171	.108583	232.93	0.000	25.07886	25.50455

BMI, Medicaid, Health Insurance, and Age in 1990
The addition of the age variable does not change the correlation of Medicaid on BMI. It is the same in this regression, with a 1.688 point coefficient correlation. Health insurance has a large change in this regression, going from the baseline of 0.018 to to -1.163. Age has a small increase in this regression, with a 0.087 point increase. The addition of the new Age coefficient could be the reason behind this increase.

. regress BMI90 YMed90 YHI90 Age90

Source	SS	df	MS		Number of obs	=	9,797
					F(3, 9793)	=	33.52
Model	2450.62937	3	816.876456		Prob > F	=	0.0000
Residual	238654.945	9,793	24.3699525		R-squared	=	0.0102
					Adj R-squared	=	0.0099
Total	241105.575	9,796	24.6126556		Root MSE	=	4.9366

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.687064	.2059031	8.19	0.000	1.283452	2.090677
YHI90	-.0163445	.1239559	-0.13	0.895	-.2593237	.2266346
Age90	.086652	.0220646	3.93	0.000	.0434009	.1299031
_cons	22.7963	.6446145	35.36	0.000	21.53272	24.05988

BMI, Medicaid, Health Insurance, Age, and Gender in 1990
The addition of the Gender Variable, Fem, makes a significant impact on BMI. Those on Medicaid have a BMI 2.120 points higher than everyone else in the sample. Being Female correlates to a -1.110 decrease in BMI in comparison to being Male. This speaks volumes to the relationship between gender and BMI. Age is statistically significant in this regression. A year increase in age correlates to a 0.094 point increase in BMI. Age continues to fluctuate with the addition of variables, going from negative in the regression to positive in this regression.

. regress BMI90 YMed90 YHI90 Age90 Fem

Source	SS	df	MS		Number of obs	=	9,797
					F(4, 9792)	=	55.62
Model	5356.50159	4	1339.1254		Prob > F	=	0.0000
Residual	235749.073	9,792	24.0756815		R-squared	=	0.0222
					Adj R-squared	=	0.0218
Total	241105.575	9,796	24.6126556		Root MSE	=	4.9067

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	2.120324	.2084212	10.17	0.000	1.711775	2.528873
YHI90	.1050032	.1236994	0.85	0.396	-.1374731	.3474795
Age90	.0943676	.0219422	4.30	0.000	.0513563	.1373788
Fem	-1.110351	.1010674	-10.99	0.000	-1.308464	-.9122378
_cons	23.02146	.6410385	35.91	0.000	21.7649	24.27803

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity in 1990
This regression indicates that an individual on Medicaid would have a BMI 1.900 points higher than those in the NLSY sample. Private health insurance in statistically significant in this sample, as is all the other variables, and provides a small BMI point increase - those with Health Insurance would have a 0.288 increase in BMI to those who are not insured. Another year of Age would correlate to a 0.100 increase in

BMI. Females in this sample have a BMI of -1.077 less than Males. Relative to White individuals, Black and Hispanic individuals have significantly larger BMIs, at 1.553 and 1.407, respectively.

```
. regress BMI90 YMed90 YHI90 Age90 Fem Black Hisp
```

Source	SS	df	MS		Number of obs	=	9,797
					F(6, 9790)	=	74.85
Model	10575.4557	6	1762.57595		Prob > F	=	0.0000
Residual	230530.119	9,790	23.5475096		R-squared	=	0.0439
					Adj R-squared	=	0.0433
Total	241105.575	9,796	24.6126556		Root MSE	=	4.8526

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.889557	.2070293	9.13	0.000	1.483736	2.295377
YHI90	.2879553	.1229897	2.34	0.019	.0468701	.5290404
Age90	.1008238	.0217056	4.65	0.000	.0582763	.1433712
Fem	-1.07675	.0999849	-10.77	0.000	-1.272741	-.8807588
Black	1.553151	.1175652	13.21	0.000	1.322699	1.783603
Hisp	1.406519	.1377428	10.21	0.000	1.136514	1.676523
_cons	22.07134	.6372721	34.63	0.000	20.82215	23.32052

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, and Income in 1990
This regression, with the added Income variable, does not make a significant difference for those on Medicaid in this sample. It also does not create a significant difference for the previously used variables: Health Insurance, Age, Gender, and Ethnicity. There is a correlation between Income and BMI - those with higher incomes in this sample have lower BMIs, indicating a downward bias.

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, and Education in 1990
This regression, with the added Education variable, does not make a significant difference for the other variables in this sample, except for the Income variable. With regard to the Education variable, Education's impact on BMI in this sample indicates that those with higher education levels have lower BMIs.

Source	SS	df	MS		Number of obs	=	9,595
					F(10, 9584)	=	45.53
Model	10720.0076	10	1072.00076		Prob > F	=	0.0000
Residual	225652.742	9,584	23.5447352		R-squared	=	0.0454
					Adj R-squared	=	0.0444
Total	236372.75	9,594	24.6375599		Root MSE	=	4.8523

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.910991	.2168908	8.81	0.000	1.485839	2.336143
YHI90	.3536653	.1317619	2.68	0.007	.0953841	.6119466
Age90	.1005237	.0221395	4.54	0.000	.0571256	.1439219
Fem	-1.133574	.105894	-10.70	0.000	-1.341148	-.9259993
Black	1.57117	.1198528	13.11	0.000	1.336233	1.806107
Hisp	1.428495	.1392246	10.26	0.000	1.155585	1.701405
OneInc90	.1556417	.1512795	1.03	0.304	-.1408981	.4521816
FifteenInc90	.0090044	.1682122	0.05	0.957	-.3207271	.3387358
TwentyfiveInc90	-.093471	.1806164	-0.52	0.605	-.4475175	.2605754
FiftyInc90	-.5528718	.3539791	-1.56	0.118	-1.246746	.1410022
_cons	22.04204	.6642657	33.18	0.000	20.73994	23.34414

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, and Marriage Status in 1990

This regression includes the newly added Region and Marriage variables. In this sample, living in an urban environment comparative to living in a rural environment would correlate to a -0.526 decrease in BMI. This could be because individuals who live in urban environments have better access to food, whereas those who live in rural areas may need to travel long distances to find healthy and fresh foods. Married individuals in this sample have a 0.694 increase in BMI. This increase in BMI could potentially be because married individuals now account for another spouse, giving them more factors to be responsible for. This change in household and personal lifestyle could cause an effect on BMI.

Source	SS	df	MS		Number of obs	=	9,246
					F(16, 9229)	=	34.51
Model	13011.9983	16	813.249892		Prob > F	=	0.0000
Residual	217466.454	9,229	23.5633821		R-squared	=	0.0565
					Adj R-squared	=	0.0548
Total	230478.452	9,245	24.9300651		Root MSE	=	4.8542

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.864092	.2208187	8.44	0.000	1.431239	2.296945
YHI90	.4768736	.1414035	3.37	0.001	.1996915	.7540557
Age90	.1075878	.022953	4.69	0.000	.0625949	.1525808
Fem	-.9939524	.1116204	-8.90	0.000	-1.212753	-.7751517
Black	1.605172	.1270089	12.64	0.000	1.356207	1.854138
Hisp	1.437203	.1462725	9.83	0.000	1.150477	1.72393
OneHS90	-.163323	.1495827	-1.09	0.275	-.4565381	.1298921
TwoHS90	-.3291742	.1750161	-1.88	0.060	-.6722445	.013896
ThreeHS90	-1.34067	.1927927	-6.95	0.000	-1.718586	-.9627533
OneInc90	.0404709	.1614757	0.25	0.802	-.2760571	.356999
FifteenInc90	.0167324	.1817442	0.09	0.927	-.3395264	.3729912
wentyfiveInc90	.1805891	.1987835	0.91	0.364	-.2090704	.5702487
FiftyInc90	.0089503	.3724816	0.02	0.981	-.721196	.7390965
Urb90	-.5260845	.1283363	-4.10	0.000	-.7776521	-.274517
YMarr90	.6943943	.1510847	4.60	0.000	.398235	.9905536
NMarr90	.7174131	.1593378	4.50	0.000	.4050758	1.02975
_cons	21.88308	.7178349	30.48	0.000	20.47596	23.29019

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, Marriage Status, Geographics, and Number of Dependents in 1990

This regression looks at how geographics play a role in BMI within the sample. Relative to the West, those who live in the Northeast correlate to having a 0.408 point increase in BMI, while those who live in the North Central and South have a BMI increase of 0.336 points and 0.287 points, respectively. Contrary to the initial hypothesis, those with more dependents correlate to having lower BMIs. The baseline for the Number of Dependents is 0. Those in our sample who care for one dependent correlate to a decrease in BMI by -0.336. Those in our sample who care for two and three dependents correlate to a decrease in BMI by -0.229 and -0.384.

Source	SS	df	MS		Number of obs	=	9,246
					F(22, 9223)	=	25.54
Model	13235.8621	22	601.630095		Prob > F	=	0.0000
Residual	217242.59	9,223	23.5544389		R-squared	=	0.0574
					Adj R-squared	=	0.0552
Total	230478.452	9,245	24.9300651		Root MSE	=	4.8533

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.855001	.2234197	8.30	0.000	1.417049	2.292954
YHI90	.4603762	.1418621	3.25	0.001	.1822952	.7384573
Age90	.1109038	.0233337	4.75	0.000	.0651645	.1566431
Fem	-.9869849	.1121673	-8.80	0.000	-1.206858	-.7671121
Black	1.601113	.1335034	11.99	0.000	1.339417	1.862809
Hisp	1.5413	.153169	10.06	0.000	1.241055	1.841545
OneHS90	-.1655466	.1503443	-1.10	0.271	-.4602547	.1291615
TwoHS90	-.3266672	.1772594	-1.84	0.065	-.6741348	.0208004
ThreeHS90	-1.37782	.1974548	-6.98	0.000	-1.764876	-.9907651
OneInc90	.0288426	.162468	0.18	0.859	-.2896307	.3473158
FifteenInc90	-.0133969	.1835343	-0.07	0.942	-.3731648	.346371
wentyfiveInc90	.1565513	.2002366	0.78	0.434	-.2359567	.5490593
FiftyInc90	.0042411	.3733476	0.01	0.991	-.7276028	.7360849
Urb90	-.5084305	.1309728	-3.88	0.000	-.7651662	-.2516948
YMarr90	.7065887	.1515343	4.66	0.000	.4095479	1.00363
NMarr90	.6663405	.1665776	4.00	0.000	.3398116	.9928694
NEast90	.4079136	.1699135	2.40	0.016	.0748456	.7409816
NCen90	.336393	.1633659	2.06	0.040	.0161598	.6566263
Sou90	.2865647	.1506344	1.90	0.057	-.008712	.5818414
OneDep90	-.0336373	.1444351	-0.23	0.816	-.316762	.2494874
TwoDep90	-.2289954	.1527594	-1.50	0.134	-.5284376	.0704468
ThreeDep90	-.0384524	.17937	-0.21	0.830	-.3900573	.3131525
_cons	21.60424	.7308602	29.56	0.000	20.17159	23.03689

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, Marriage Status, Geographics, Number of Dependents, Depression, Occupation, Family Size, and Weeks of Unemployment in 1990

In this regression, Depression plays an interesting role. Using the 7 point scale, those who fall under the category of OneDepre correlate to having a decrease in BMI by -0.082 points. However, as depression worsens, those who fall under the category of FiveDepre and TenDepre correlate to having a BMI increase by 0.471 points and 0.804 points, respectively. The Occupation variable indicates that those who having have more than one job have lower BMI, however there is no significant trend. Family Size and Weeks of Unemployment doesn't display any significant trends in its correlation to BMI.

Source	SS	df	MS		Number of obs	=	6,292
					F(37, 6254)	=	11.57
Model	10243.3435	37	276.847122		Prob > F	=	0.0000
Residual	149643.196	6,254	23.9275977		R-squared	=	0.0641
					Adj R-squared	=	0.0585
Total	159886.539	6,291	25.4151231		Root MSE	=	4.8916

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	2.131236	.2860628	7.45	0.000	1.570454	2.692017
YHI90	.7230762	.179175	4.04	0.000	.3718317	1.074321
Age90	.1173719	.0290935	4.03	0.000	.0603387	.174405
Fem	-.8784497	.1388448	-6.33	0.000	-1.150633	-.6062663
Black	1.372105	.162934	8.42	0.000	1.052698	1.691511
Hisp	1.349866	.1846326	7.31	0.000	.9879228	1.711809
OneHS90	-.3867904	.191781	-2.02	0.044	-.7627471	-.0108338
TwoHS90	-.5306231	.2203244	-2.41	0.016	-.9625347	-.0987116
ThreeHS90	-1.520099	.2504747	-6.07	0.000	-2.011115	-1.029082
OneInc90	.4583937	.2938255	1.56	0.119	-.1176052	1.034393
FifteenInc90	.4709627	.3106404	1.52	0.130	-.1379991	1.079925
TwentyfiveInc90	.6154265	.3242435	1.90	0.058	-.0202021	1.251055
FiftyInc90	.4807082	.5165903	0.93	0.352	-.5319861	1.493403
Urb90	-.6135017	.1653327	-3.71	0.000	-.9376106	-.2893927
YMarr90	.4991853	.2087385	2.39	0.017	.0899862	.9083845
NMarr90	.6877354	.2090183	3.29	0.001	.2779878	1.097483
NEast90	.3374128	.2097428	1.61	0.108	-.0737551	.7485807
NCen90	.2842348	.1991898	1.43	0.154	-.1062457	.6747153
Sou90	.3666746	.1844359	1.99	0.047	.005117	.7282322
OneDep90	-.2436798	.1905267	-1.28	0.201	-.6171777	.129818
TwoDep90	-.5166405	.2215097	-2.33	0.020	-.9508756	-.0824054
ThreeDep90	-.298322	.2615161	-1.14	0.254	-.8109833	.2143393
OneDepre	-.0981298	.153873	-0.64	0.524	-.3997737	.203514
FiveDepre	.3945744	.1836697	2.15	0.032	.0345188	.7546301
TenDepre	.7908002	.2225451	3.55	0.000	.3545354	1.227065
OneOcc90	-.5497866	.3261427	-1.69	0.092	-1.189138	.0895652
TwoOcc90	-.4491363	.3525856	-1.27	0.203	-1.140325	.2420525
ThreeOcc90	-.488675	.3930517	-1.24	0.214	-1.259191	.2818412
TwoFam90	.1605845	.2239212	0.72	0.473	-.278378	.599547
ThreeFam90	.8912572	.2343004	3.80	0.000	.431948	1.350567
FourFam90	.5217872	.2601371	2.01	0.045	.0118292	1.031745
FiveFam90	.5976247	.2771546	2.16	0.031	.0543065	1.140943
OneWksUnInt90	.0164668	.2168051	0.08	0.939	-.4085456	.4414792
TenWksUnInt90	.352958	.3438869	1.03	0.305	-.3211785	1.027094
TwentyWksUnInt90	.1126628	.4646918	0.24	0.808	-.7982927	1.023618
ThirtyWksUnInt90	-1.670733	1.639164	-1.02	0.308	-4.884057	1.542591
FortyWksUnInt90	-.0712519	.4968036	-0.14	0.886	-1.045158	.9026537
_cons	21.34527	.937319	22.77	0.000	19.50781	23.18274

Baseline - BMI, Medicaid, and Health Insurance in 2010

This regression indicates that an individual on Medicaid in 1990 would have a BMI 1.924 points higher than someone who was not on Medicaid. Medicaid's effect on BMI is statistically significant. Private health insurance is not statistically significant.

```
. regress BMI10 YMed10 YHI10
```

Source	SS	df	MS		
Model	1737.94829	2	868.974147		
Residual	252649.811	6,839	36.9425078		
Total	254387.759	6,841	37.1857563		

				Number of obs	=	6,842
				F(2, 6839)	=	23.52
				Prob > F	=	0.0000
				R-squared	=	0.0068
				Adj R-squared	=	0.0065
				Root MSE	=	6.078

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.924323	.3011708	6.39	0.000	1.333934	2.514711
YHI10	.1750724	.1869947	0.94	0.349	-.1914953	.5416401
_cons	28.76926	.1656688	173.66	0.000	28.44449	29.09402

BMI, Medicaid, Health Insurance, Age, and Gender in 2010

The addition of the Gender Variable, Fem, makes a significant impact on BMI. Being Female correlates to a -0.074 decrease in BMI in comparison to being Male. Age is not statistically significant in this regression. A year increase in age correlates to a 0.006 point increase in BMI.

Source	SS	df	MS		
Model	1748.49989	4	437.124973		
Residual	252639.259	6,837	36.9517711		
Total	254387.759	6,841	37.1857563		

				Number of obs	=	6,842
				F(4, 6837)	=	11.83
				Prob > F	=	0.0000
				R-squared	=	0.0069
				Adj R-squared	=	0.0063
				Root MSE	=	6.0788

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.93153	.3017385	6.40	0.000	1.340029	2.523032
YHI10	.1790934	.1874554	0.96	0.339	-.1883774	.5465642
Age10	.0062632	.0329665	0.19	0.849	-.0583614	.0708878
Fem	-.0738306	.1473384	-0.50	0.616	-.3626597	.2149984
_cons	28.49932	1.604866	17.76	0.000	25.35328	31.64535

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity in 2010

This regression indicates that an individual on Medicaid would have a BMI 1.688 points higher than those in the NLSY sample. Private health insurance in statistically significant in this sample, as is all the other variables, and provides a small BMI point increase - those with Health Insurance would have a 0.514 increase in BMI to those who are not insured. Another year of Age would correlate to a 0.005 increase in BMI. Females in this sample have a BMI of -0.086 less than Males. Relative to White individuals, Black and Hispanic individuals have significantly larger BMIs, at 2.017 and 1.519, respectively.

In comparison to the regression ran on the 1990 variables, individuals who are Black and Hispanic have a higher correlation to an increase in BMI.

Source	SS	df	MS		Number of obs	=	6,842
					$F(6, 6835)$	=	34.08
Model	7389.96906	6	1231.66151		Prob > F	=	0.0000
Residual	246997.79	6,835	36.1372041		R-squared	=	0.0291
					Adj R-squared	=	0.0282
Total	254387.759	6,841	37.1857563		Root MSE	=	6.0114

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.688327	.2995854	5.64	0.000	1.101047	2.275608
YHI10	.514805	.1874594	2.75	0.006	.1473262	.8822838
Age10	.005352	.0326027	0.16	0.870	-.0585594	.0692634
Fem	-.0858311	.1457157	-0.59	0.556	-.3714792	.199817
Black	2.017423	.1702406	11.85	0.000	1.683698	2.351147
Hisp	1.518982	.1968517	7.72	0.000	1.133092	1.904873
_cons	27.42355	1.589787	17.25	0.000	24.30707	30.54002

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, and Income in 2010
This regression, with the added Income variable, does not make a significant difference for those on Medicaid in this sample. It also does not create a significant difference for the previously used variables: Health Insurance, Age, Gender, and Ethnicity. Unlike in the 1990 regression of the same variables, Income does not show a significant trend. OneInc correlates to a decrease in BMI points, while attaining more income then correlates to an increase in BMI points.

Source	SS	df	MS		Number of obs	=	5,657
					$F(10, 5646)$	=	16.63
Model	5732.9641	10	573.29641		Prob > F	=	0.0000
Residual	194691.988	5,646	34.4831717		R-squared	=	0.0286
					Adj R-squared	=	0.0269
Total	200424.952	5,656	35.4358118		Root MSE	=	5.8722

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.492495	.4054998	3.68	0.000	.6975596	2.28743
YHI10	.4638936	.2383346	1.95	0.052	-.0033339	.9311211
Age10	-.021507	.0351378	-0.61	0.541	-.0903906	.0473765
Fem	-.1546783	.1638882	-0.94	0.345	-.4759623	.1666056
Black	2.108645	.1855295	11.37	0.000	1.744936	2.472354
Hisp	1.415754	.2102083	6.74	0.000	1.003665	1.827843
OneInc10	-.1233266	.3623177	-0.34	0.734	-.8336085	.5869553
FifteenInc10	.4339922	.3564886	1.22	0.223	-.2648624	1.132847
Twentyfive~10	.3303399	.3299172	1.00	0.317	-.3164246	.9771043
FiftyInc10	.1700286	.3418583	0.50	0.619	-.5001449	.8402022
_cons	28.64469	1.740612	16.46	0.000	25.23242	32.05696

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, and Education in 2010
This regression, with the added Education variable, does not make a significant difference for the other variables in this sample, except for the Income variable. In the previous regression, the added Income variable shows a decrease in BMI correlated with higher income levels. However, Income is not statistically significant at the 95% level. With regard to the Education variable, Education's impact on BMI in this sample indicates that those with education levels have lower BMIs. The same trend is prevalent in 1990.

Source	SS	df	MS		Number of obs	=	5,588
					F(13, 5574)	=	13.04
Model	5817.85863	13	447.527587		Prob > F	=	0.0000
Residual	191267.736	5,574	34.3142691		R-squared	=	0.0295
					Adj R-squared	=	0.0273
Total	197085.594	5,587	35.2757463		Root MSE	=	5.8578

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.505435	.4073914	3.70	0.000	.706789	2.304081
YHI10	.467052	.2402671	1.94	0.052	-.0039652	.9380691
Age10	-.0270478	.0353139	-0.77	0.444	-.0962767	.0421812
Fem	-.1593473	.1652304	-0.96	0.335	-.4832633	.1645688
Black	2.020174	.186589	10.83	0.000	1.654386	2.385961
Hisp	1.395123	.2121994	6.57	0.000	.9791299	1.811117
OneInc10	-.1551042	.3646679	-0.43	0.671	-.8699954	.559787
FifteenInc10	.3596232	.3590775	1.00	0.317	-.3443086	1.063555
Twentyfive~10	.2704061	.3326151	0.81	0.416	-.3816491	.9224613
FiftyInc10	.2017103	.3463499	0.58	0.560	-.4772703	.880691
OneHS10	.4990452	.2049589	2.43	0.015	.0972459	.9008446
TwoHS10	.2780074	.3021816	0.92	0.358	-.3143864	.8704011
ThreeHS10	-.1051296	.2690852	-0.39	0.696	-.6326416	.4223823
_cons	28.68695	1.75642	16.33	0.000	25.24368	32.13022

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, and Marriage Status in 2010

This regression includes the newly added Region and Marriage variables. In comparison to the 1990 regression, another year of age now correlates to a decrease in BMI points, as opposed to an increase. In this sample, living in an urban environment comparative to living in a rural environment would correlate to a -0.391 decrease in BMI. Married individuals in this sample have a 0.443 point increase in BMI.

Source	SS	df	MS		Number of obs	=	5,552
					F(16, 5535)	=	11.53
Model	6324.16178	16	395.260111		Prob > F	=	0.0000
Residual	189793.855	5,535	34.289766		R-squared	=	0.0322
					Adj R-squared	=	0.0294
Total	196118.017	5,551	35.3302138		Root MSE	=	5.8557

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.53563	.4094767	3.75	0.000	.7328947	2.338365
YHI10	.4196436	.2466378	1.70	0.089	-.0638632	.9031505
Age10	-.0283064	.0354169	-0.80	0.424	-.0977374	.0411245
Fem	-.1143606	.16669	-0.69	0.493	-.4411384	.2124173
Black	2.064278	.1945242	10.61	0.000	1.682934	2.445621
Hisp	1.497166	.2171444	6.89	0.000	1.071477	1.922854
OneInc10	-.1187738	.3669853	-0.32	0.746	-.838209	.6006614
FifteenInc10	.4193059	.3606095	1.16	0.245	-.2876303	1.126242
Twentyfive~10	.3212969	.3342745	0.96	0.337	-.3340124	.9766063
FiftyInc10	.265919	.3484547	0.76	0.445	-.4171891	.9490271
OneHS10	.4891203	.2059551	2.37	0.018	.0853675	.8928731
TwoHS10	.2609684	.3040127	0.86	0.391	-.3350158	.8569527
ThreeHS10	-.1245887	.2704921	-0.46	0.645	-.6548593	.405682
Urb10	-.3911569	.1769323	-2.21	0.027	-.7380136	-.0443002
YMarr10	.4427237	.1913773	2.31	0.021	.067549	.8178985
NMarr10	.7382843	.2533193	2.91	0.004	.2416789	1.23489
_cons	28.60044	1.766817	16.19	0.000	25.13679	32.0641

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, Marriage Status, Geographics, and Number of Dependents in 2010
The addition of the Geographics variable in this regression does not seem to portray any significant differences in comparison to the 1990 regression. Similar to the 1990 regression results on the Geographics variable, living in the Northeast, North Central, or South all would correlate to a positive increase in BMI, relative to living in the West.

Similar to the 1990 regression, those with more dependents correlate to having lower BMIs in the 2010 regression as well. The baseline for the Number of Dependents is 0. Those in our sample who care for one dependent correlate to a decrease in BMI by -0.693. Those in our sample who care for two and three dependents correlate to a decrease in BMI by -0.371 and -0.168. Having more dependents in 2010 correlates to an even larger decrease in BMI points than it did in 1990. This could be due to the fact that it is more difficult and strenuous for individuals to care for self with dependents.

Source	SS	df	MS			
Model	6835.76764	22	310.716711			
Residual	189282.249	5,529	34.2344455			
Total	196118.017	5,551	35.3302138			

	Number of obs	=	5,552
F(22, 5529)	=	9.08
Prob > F	=	0.0000
R-squared	=	0.0349
Adj R-squared	=	0.0310
Root MSE	=	5.851

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.600063	.4106739	3.90	0.000	.7949812	2.405146
YHI10	.4268773	.2472487	1.73	0.084	-.0578275	.911582
Age10	-.026458	.0354207	-0.75	0.455	-.0958964	.0429804
Fem	-.1169746	.166978	-0.70	0.484	-.4443172	.210368
Black	1.94259	.2055014	9.45	0.000	1.539727	2.345454
Hisp	1.55529	.2299659	6.76	0.000	1.104466	2.006114
OneInc10	-.1010074	.3686747	-0.27	0.784	-.8237548	.62174
FifteenInc10	.4112717	.3614578	1.14	0.255	-.2973277	1.119871
Twentyfive~10	.3491278	.3358285	1.04	0.299	-.3092281	1.007484
FiftyInc10	.3072021	.3501636	0.88	0.380	-.3792563	.9936604
OneHS10	.5140905	.2060925	2.49	0.013	.1100681	.9181129
TwoHS10	.293153	.3040768	0.96	0.335	-.3029571	.8892631
ThreeHS10	-.1307648	.2708909	-0.48	0.629	-.6618175	.4002879
Urb10	-.3046779	.1797952	-1.69	0.090	-.6571471	.0477914
YMarr10	.462185	.1923536	2.40	0.016	.0850964	.8392736
NMarr10	.6794909	.268727	2.53	0.011	.1526803	1.206302
NEast10	.1679435	.276016	0.61	0.543	-.3731563	.7090432
NCen10	.3396375	.2543308	1.34	0.182	-.1589508	.8382259
Sou10	.6036947	.2313089	2.61	0.009	.1502384	1.057151
OneDep10	-.6926912	.2762228	-2.51	0.012	-1.234197	-.1511859
TwoDep10	-.3709147	.2469431	-1.50	0.133	-.8550203	.1131908
ThreeDep10	-.1676113	.2517431	-0.67	0.506	-.6611267	.3259041
_cons	28.35793	1.791234	15.83	0.000	24.8464	31.86945

BMI, Medicaid, Health Insurance, Age, Gender, Ethnicity, Income, Education, Region, Marriage Status, Geographics, Number of Dependents, Depression, Occupation, Family Size, and Weeks of Unemployment in 2010
In this regression, and similar to the 1990 regression, Depression plays an interesting role. Using the 7 point scale, those who fall under the category of OneDepre correlate to having a decrease in BMI by -0.006 points. However, as depression worsens, those who fall under the category of FiveDepre and TenDepre correlate to having a BMI increase of 0.929 points and 1.41 points, respectively.

The Occupation variable indicates that those who having have more than one job have lower BMI, however there is no significant trend. Family Size and Weeks of Unemployment doesn't display any notable trends in its correlation to BMI.

Source	SS	df	MS		Number of obs	=	5,157
					F(37, 5119)	=	6.46
Model	8033.5884	37	217.124011		Prob > F	=	0.0000
Residual	172080.143	5,119	33.6159685		R-squared	=	0.0446
					Adj R-squared	=	0.0377
Total	180113.731	5,156	34.9328416		Root MSE	=	5.7979

BMI10	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.305957	.4551458	2.87	0.004	.4136767	2.198238
YHI10	.6654848	.2614496	2.55	0.011	.1529318	1.178038
Age10	-.0619751	.037052	-1.67	0.094	-.1346129	.0106627
Fem	-.289855	.1757818	-1.65	0.099	-.6344624	.0547524
Black	1.931195	.2122377	9.10	0.000	1.515118	2.347271
Hisp	1.449137	.2379087	6.09	0.000	.9827346	1.91554
OneInc10	.227679	.4536404	0.50	0.616	-.6616501	1.117008
FifteenInc10	.8394797	.448758	1.87	0.061	-.0402779	1.719237
TwentyfiveInc10	.6748144	.427288	1.58	0.114	-.1628527	1.512482
FiftyInc10	.7374399	.4390924	1.68	0.093	-.123369	1.598249
OneHS10	.5691118	.2125476	2.68	0.007	.1524276	.985796
TwoHS10	.3051515	.3134792	0.97	0.330	-.3094018	.9197047
ThreeHS10	-.0360219	.2780918	-0.13	0.897	-.5812007	.5091569
Urb10	-.2687526	.1852057	-1.45	0.147	-.631835	.0943298
YMarr10	.2989047	.241105	1.24	0.215	-.1737641	.7715735
NMarr10	.7157849	.279732	2.56	0.011	.1673906	1.264179
NEast10	.1463145	.2846407	0.51	0.607	-.4117029	.7043319
NCen10	.3569868	.2622526	1.36	0.173	-.1571403	.871114
Sou10	.6525898	.2386691	2.73	0.006	.1846963	1.120483
OneDep10	-.5867902	.291992	-2.01	0.045	-1.159219	-.0143611
TwoDep10	-.2294852	.2701211	-0.85	0.396	-.7590381	.3000677
ThreeDep10	-.0678797	.2823009	-0.24	0.810	-.6213102	.4855507
OneDepre	-.0298807	.1960796	-0.15	0.879	-.4142806	.3545192
FiveDepre	.9090421	.2402988	3.78	0.000	.4379537	1.380131
TenDepre	1.435444	.3183734	4.51	0.000	.8112956	2.059592
OneOcc10	-.6324076	.4971426	-1.27	0.203	-1.60702	.3422045
TwoOcc10	-.4548849	.5250242	-0.87	0.386	-1.484157	.574387
ThreeOcc10	-.335005	.6146745	-0.55	0.586	-1.54003	.8700198
TwoFam10	.5494614	.2722067	2.02	0.044	.01582	1.083103
ThreeFam10	.2158864	.3029672	0.71	0.476	-.378059	.8098317
FourFam10	.1011629	.3372669	0.30	0.764	-.5600244	.7623502
FiveFam10	.5413029	.3762481	1.44	0.150	-.1963042	1.27891
OneWksUnInt10	.2452741	.3877607	0.63	0.527	-.5149027	1.005451
TenWksUnInt10	.0659743	.5119033	0.13	0.897	-.937575	1.069524
TwentyWksUnInt10	-.3808239	.5907175	-0.64	0.519	-1.538883	.777235
ThirtyWksUnInt10	-.2438467	.6723705	-0.36	0.717	-1.56198	1.074287
FortyWksUnInt10	.1155923	.4650882	0.25	0.804	-.7961794	1.027364
_cons	29.55225	1.887976	15.65	0.000	25.85101	33.25349

Special Regressions (shows income and education variable are correlated)
Income in 1990
In this regression, Income is singled out. There is a clear indication that an increase in Income correlates to a decrease in BMI points. This trend continues, even with the addition of the Medicaid, Health

Insurance, Age, Gender, and Ethnicity variables. With all of the variables included, Income's correlation to BMI's decrease continues, but at a smaller scale.

Education in 1990
In this regression of the Education variable alone, there is a clear indication that an increase in Education correlates to a decrease in BMI points. This trend continues, even with the addition of the Medicaid, health Insurance, Age, Gender, and Ethnicity variables. With all of the variables included, Education correlation to BMI's decrease continues.

Income in 2010
In this regression of the Income variable alone in 2010, there is an interesting trend with income. The -0.159 coefficient on OneInc10 indicates that those who make under $15,000 correlate to a decrease in BMI points. However, those who fall into the $15,000 - $25,000 income bracket correlate to a 0.411 point increase in BMI. The higher the income level from that point correlates to a gradual decline in BMI. Unlike in the 1990 Income regression, there is no definite and consistent trend. The addition of new variables doesn't change this trend.

Education in 2010
In this regression of the Education variable alone, there is a clear indication that an increase in Education correlates to a decrease in BMI points. This trend continues, even with the addition of the Medicaid, health Insurance, Age, Gender, and Ethnicity variables. With all of the variables included, Education correlation to BMI's decrease continues.

Interaction Terms in 1990 - Females on Medicaid and Private Health Insurance
In this regression, the interaction terms are Females on Medicaid and Females on Health Insurance. Females on Medicaid in this study correlate to having a 1.406 increase in BMI in comparison to Males with no health insurance. Females on private health insurance in this study correlate to having a -0.578 decrease in BMI in comparison to Males with no insurance
. With the other added variables, this trend remains constant.

Interaction Terms in 2010
In 2010, Females on Medicaid in this study correlate to having a 0.407 point increase in BMI in comparison to Males with no health insurance. Females on private health insurance correlate to having an even larger decrease in BMI points, at -1.739.

No Gender in 1990
Gender plays a large role in the coefficient change on the variables. In 1990, when Gender is removed from the regression, the coefficient on Medicaid drops from 1.911 to 1.501. The Gender variable has the largest effect on the various health insurance-related variables, and a slight effect on the Income variable. According to Miltra Toosi's "A century of change: the U.S. labor force, 1950–2050", women in the labor force increased at an extremely rapid pace in the last 50 years. The workforce is increasingly made up of more and more women, leading to an increase in employer provided health care. When looking at the regression for the Male and private Health Insurance in 1990 and 2010, there is an approximately 1,600 observation decrease in the number of males who have private health insurance from 1990 to 2010. When the Male variable is removed, the coefficient decreases, indicating a downward bias. This indicates a macroeconomic phenomenon of more women having private health insurance

It does not make a significant difference on the ethnicity variables.

In the 2010 regression, gender does not play a role in the coefficient change on any of the variables, except for the Depression variable. In 20 years, women have played a larger role within the labor force, potentially explaining why there is an effect in 1990 and not in 2010.

Consistent in 1990

There is an approximately 5,000 observation drop going from the 1990 regression to the 2010 regression. When the regression is made consistent by taking the number of surveyed individuals from 2010 and only observing them in the 1990 regression, the coefficient on Medicaid increases from 1.687 in the all-inclusive 1990 data to 2.329 in the uniformed data.

Looking at the summary statistics indicates that all of the sub variables nearly double from the consistent data to the all-inclusive data.

```
. regress BMI90 YMed90 YHI90 Age90 Fem Black Hisp
```

Source	SS	df	MS		Number of obs	=	9,797
					F(6, 9790)	=	74.85
Model	10575.4557	6	1762.57595		Prob > F	=	0.0000
Residual	230530.119	9,790	23.5475096		R-squared	=	0.0439
					Adj R-squared	=	0.0433
Total	241105.575	9,796	24.6126556		Root MSE	=	4.8526

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.889557	.2070293	9.13	0.000	1.483736	2.295377
YHI90	.2879553	.1229897	2.34	0.019	.0468701	.5290404
Age90	.1008238	.0217056	4.65	0.000	.0582763	.1433712
Fem	-1.07675	.0999849	-10.77	0.000	-1.272741	-.8807588
Black	1.553151	.1175652	13.21	0.000	1.322699	1.783603
Hisp	1.406519	.1377428	10.21	0.000	1.136514	1.676523
_cons	22.07134	.6372721	34.63	0.000	20.82215	23.32052

With consistent sample

```
. regress BMI90 YMed90 YHI90 Age90 Fem Black Hisp if consample==1
```

Source	SS	df	MS		Number of obs	=	4,853
					F(6, 4846)	=	37.60
Model	5029.94952	6	838.32492		Prob > F	=	0.0000
Residual	108049.368	4,846	22.2966092		R-squared	=	0.0445
					Adj R-squared	=	0.0433
Total	113079.318	4,852	23.3057127		Root MSE	=	4.7219

BMI90	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	2.416925	.2958137	8.17	0.000	1.836996	2.996854
YHI90	.53338	.1821995	2.93	0.003	.1761864	.8905736
Age90	.0861334	.0304804	2.83	0.005	.0263779	.1458889
Fem	-.9942226	.1384042	-7.18	0.000	-1.265558	-.7228876
Black	1.411102	.162114	8.70	0.000	1.093285	1.728919
Hisp	1.224364	.1843387	6.64	0.000	.8629761	1.585751
_cons	22.3422	.8996608	24.83	0.000	20.57846	24.10595

BMI and Medicaid in 1990

This regression indicates that those on Medicaid have a BMI 1.684 points higher than everyone else in the sample. The NLS is a nationally representative sample that over samples low-income individuals and minorities, possibly explaining the correlation between Medicaid and BMI. Low-income individuals are financially constrained, potentially giving them less options to purchase and consume healthy foods. According to the State of Obesity, children living below the federal household poverty level have an obesity rate 2.7 times higher (27.4 percent) than children living in households exceeding 400 percent of the federal poverty level.

```
. regress BMI90 YMed90, robust

Linear regression                               Number of obs   =       9,797
                                                F(1, 9795)      =       53.89
                                                Prob > F        =      0.0000
                                                R-squared       =      0.0086
                                                Root MSE        =        4.94
```

BMI90	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.684134	.2294252	7.34	0.000	1.234413	2.133854
_cons	25.3055	.0505897	500.21	0.000	25.20633	25.40466

BMI and Medicaid in 2010
By 2010, those on Medicaid had BMIs 1.797 points higher than others in the NLSY. This difference is larger than in 1990, however, this could be attributed to the lower number of observations, the lower amount of people surveyed, and/or the pervasiveness of obesity. This observed decrease could have been due to less people being surveyed-- the surveyed individuals could have moved, changed their contact information, passed away, or became ineligible to partake in the survey. In 2010, approximately 3,000 fewer individuals were surveyed.

```
. regress BMI10 YMed10, robust

Linear regression                               Number of obs   =       6,842
                                                F(1, 6840)      =       29.14
                                                Prob > F        =      0.0000
                                                R-squared       =      0.0067
                                                Root MSE        =       6.078
```

BMI10	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.786906	.3309977	5.40	0.000	1.138047	2.435764
_cons	28.90667	.0744864	388.08	0.000	28.76066	29.05269

BMI, Medicaid, and Health Insurance in 1990
This regression indicates that those on Medicaid have BMI 1.698 higher than those in the sample. The addition of the private Health Insurance variable decreases BMI by 0.01, which is significant. Since it is unlikely that an adult individual's height increases within one year, the individual's weight must have increased. In comparison to Health Insurance, an individual on Medicaid has a significantly higher BMI. Medicaid is tailored towards lower-income individuals, whereas private health insurance is oftentimes costly. This income differentially potentially explains this disparity.

The BMI amongst individuals who have Medicaid continues to increase with the addition of new variables because of the omitted variable bias. Given the multitude of omitted variables, the coefficients should not be interpreted as the causal effect of Medicaid and Health Insurance on BMI. In this regression, No Health Insurance is also included. The reference group is No Health Insurance because of multicollinearity. An extra insurance variable is not included in this regression because of a missing category. Medicaid is statistically significant at the 95% level; however, Health Insurance and No Health Insurance aren't. Despite not being statistically significant, the aforementioned variables are still increasing. In addition, the standard error on the Medicaid and private health insurance variables rise as well.

```
. regress BMI90 YMed90 YHI90, robust

Linear regression                              Number of obs   =      9,797
                                               F(2, 9794)      =      26.95
                                               Prob > F        =     0.0000
                                               R-squared       =     0.0086
                                               Root MSE        =     4.9402
```

BMI90	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.697922	.2484756	6.83	0.000	1.210858	2.184985
YHI90	.0179061	.1222099	0.15	0.884	-.2216504	.2574626
_cons	25.29171	.107974	234.24	0.000	25.08006	25.50336

BMI, Medicaid, and Health Insurance in 2010
The BMI coefficients for Medicaid and Health Insurance both increase from 1990 to 2010 by approximately 0.3.
With the addition of a new private Health Insurance variable, Medicaid's effect on BMI increases for 1990 and increases for 2010. Uninsured is not in the included category because of perfect collinearity, however it is the baseline for which the other health insurance variables are compared to.

In comparison to 1990, all the coefficients decrease by approximately 1.0, with No Health Insurance having the greatest impact. All of the variables are all statistically significant at the 95% level. With the additional variable (No Health Insurance) added, the coefficient on all the other variables also decrease.

```
. regress BMI10 YMed10 YHI10, robust

Linear regression                                Number of obs   =      6,842
                                                 F(2, 6839)      =      14.99
                                                 Prob > F        =     0.0000
                                                 R-squared       =     0.0068
                                                 Root MSE        =      6.078

                         Robust
      BMI10      Coef.   Std. Err.      t     P>|t|     [95% Conf. Interval]

     YMed10   1.924323   .3619666     5.32    0.000     1.214756     2.63389
      YHI10   .1750724   .1843148     0.95    0.342    -.1862419    .5363867
      _cons   28.76926   .1642966   175.11    0.000     28.44718    29.09133
```

BMI, Medicaid, Health Insurance, and Age in 1990

In the regression of BMI and Medicaid, having Medicaid versus not having it correlates to an increase in an individual's BMI by 1.1923. Multicollinearity is present here, as not having Medicaid will decrease an individual's BMI by the same amount. Having private Health Insurance in 1990 correlates to an increase in an individual's BMI and is not statistically significant. Its t-statistic stays the same with the addition of age, conveying that age does not have a large impact.

As an individual ages by one year, their BMI increases by 0.087. The t-statistic for both the Medicaid is above 1.96, indicating that it is statistically significant at a 95% level.

```
. regress BMI90 YMed90 YHI90 Age90, robust

Linear regression                                Number of obs   =      9,797
                                                 F(3, 9793)      =      22.83
                                                 Prob > F        =     0.0000
                                                 R-squared       =     0.0102
                                                 Root MSE        =     4.9366

                         Robust
      BMI90      Coef.   Std. Err.      t     P>|t|     [95% Conf. Interval]

     YMed90   1.687064   .2481073     6.80    0.000     1.200723     2.173406
      YHI90  -.0163445    .122624    -0.13    0.894    -.2567129     .2240238
      Age90   .086652    .0217645     3.98    0.000     .0439891     .1293149
      _cons   22.7963    .6323484    36.05    0.000     21.55676     24.03583
```

BMI, Medicaid and Age in 2010

In the regression of BMI and Medicaid, having Medicaid correlates to an increase in BMI by 1.923. There is a marginal increase in BMI given a lower population.

As an individual ages by one year, their BMI increases by 0.006. The t-statistic for both the Medicaid is above 1.96 but is not for the Age variable. However, the t-statistic is small. In 2010, the marginal increase in BMI given Medicaid and Age is approximately the same 1990.

```
. regress BMI10 YMed10 YHI10 Age10, robust

Linear regression                                    Number of obs   =      6,842
                                                     F(3, 6838)      =      10.03
                                                     Prob > F        =     0.0000
                                                     R-squared       =     0.0068
                                                     Root MSE        =     6.0785
```

BMI10	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.923151	.3621784	5.31	0.000	1.213168	2.633133
YHI10	.1737562	.1846441	0.94	0.347	-.1882036	.535716
Age10	.0061189	.0324925	0.19	0.851	-.0575765	.0698143
_cons	28.47312	1.576204	18.06	0.000	25.38327	31.56297

BMI, Medicaid, Health Insurance, Age, Ethnicity, and Gender in 1990

In this regression of BMI, gender and ethnicity are added. White and Male are the baseline for comparison in this regression. Medicaid correlates to an increase in BMI by 1.890 and Health Insurance correlates to an increase in BMI by 0.288. Age is relevant in this regression, with a coefficient of 0.100. Females have a lower BMI in comparison to Males by -1.077. Black and Hispanics both have a positive correlation in BMI in comparison to White individuals. All of the t-statistics in this regression are statistically significant at the 95% level.

Jeffrey Kling and Jeffrey Liebman's research, documented within the National Bureau of Economic Research helps to explain the positive Male coefficient (2004). Their study, which examines the effects of moving youth out of high-poverty neighborhoods on overall health, presented secondary results that affected the gender domain. Females benefited from being in the experimental group - in addition to better health outcomes, they had improvements in mental health and education, and were less likely to engage in risky behavior. While this study did not look at BMI directly, its effects on females can help to explain the male versus female health outcome dichotomy.

Compared to 1990, with the addition of the new variables, the f-statistic increased significantly. With an F-stat of 70.80, the null hypothesis that all the coefficients jointly equal 0 can easily be rejected.

Looking at the robust standard errors shows that there is a correlation between those who have Medicaid and the Hispanic Black and Female variables.

```
. regress BMI90 YMed90 YHI90 Age90 Hisp Black Fem, robust

Linear regression                                    Number of obs   =      9,797
                                                     F(6, 9790)      =      70.80
                                                     Prob > F        =     0.0000
                                                     R-squared       =     0.0439
                                                     Root MSE        =     4.8526
```

BMI90	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.889557	.2507706	7.53	0.000	1.397994	2.381119
YHI90	.2879553	.1227104	2.35	0.019	.0474176	.528493
Age90	.1008238	.0214336	4.70	0.000	.0588095	.142838
Hisp	1.406519	.1388419	10.13	0.000	1.13436	1.678677
Black	1.553151	.1233926	12.59	0.000	1.311276	1.795026
Fem	-1.07675	.0986173	-10.92	0.000	-1.27006	-.8834396
_cons	22.07134	.6246015	35.34	0.000	20.84699	23.29569

```
. corr YMed90 Black Hisp Fem
(obs=10,182)
```

	YMed90	Black	Hisp	Fem
YMed90	1.0000			
Black	0.1275	1.0000		
Hisp	0.0195	-0.2640	1.0000	
Fem	0.1690	-0.0107	-0.0061	1.0000

BMI, Medicaid, Health Insurance, Age, Ethnicity, and Gender in 2010

In the regression of BM in 2010, Age and Female are not statistically significant in comparison to the 1990 regression. All of the variable coefficients are positive, excluding Female.

Going from 1990 to 2010, the R-squared value goes from explaining less than 1% of the variation to approximately 3% of the variation in BMI. There isn't a large change in robust standard error for the Medicaid and Health Insurance variables.

```
. regress BMI10 YMed10 YHI10 Age10 Hisp Black Fem, robust

Linear regression                              Number of obs    =      6,842
                                               F(6, 6835)       =      31.03
                                               Prob > F         =     0.0000
                                               R-squared        =     0.0291
                                               Root MSE         =     6.0114
```

BMI10	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.688327	.3625264	4.66	0.000	.9776628	2.398992
YHI10	.514805	.186019	2.77	0.006	.1501499	.8794602
Age10	.005352	.0321483	0.17	0.868	-.0576687	.0683728
Hisp	1.518982	.1889202	8.04	0.000	1.14864	1.889325
Black	2.017423	.1752124	11.51	0.000	1.673952	2.360893
Fem	-.0858311	.144897	-0.59	0.554	-.3698742	.198212
_cons	27.42355	1.555105	17.63	0.000	24.37506	30.47204

```
. corr YMed10 Black Hisp Fem
(obs=7,048)
```

	YMed10	Black	Hisp	Fem
YMed10	1.0000			
Black	0.1458	1.0000		
Hisp	0.0044	-0.3216	1.0000	
Fem	0.0306	0.0033	0.0030	1.0000

BMI, Medicaid, Age, Gender, Ethnicity, Income, and Education in 1990
In this regression, No HS and and 0 are the baseline for the Education and Income variable. With the addition of the new variables, Medicaid correlates to an increase in BMI by 1.841, while Health Insurance correlates to an increase in BMI by 0.504. As education level increases, BMI decreases. Nearly 33 percent of adults who did not graduate high school were obese, compared with the 21.5 percent of those who graduated from college, according to the State of Obesity. This explains the relationship between Education and BMI. There is no pattern between BMI and Income.

With the addition of new variables, the Medicaid coefficient decreases while still being statistically significant. The Education variable contributes to this decrease. An extra year of Age correlates to an increase in BMI by 0.010, while Female BMI decreases by -1.047 in comparison to Male. Hispanic and Black are both statistically significant and have a positive increase in BMI.

BMI decreases if an individual has some high school education and continues to decrease as an individual goes through more schooling. Ultimately, a college graduate has the lowest BMI in comparison to someone who only went through or completed junior high, with a -1.322 coefficient.

```
regress BMI90 YMed90 YHI90 Age90 Hisp Black Fem OneHS90 TwoHS90 ThreeHS90 OneInc90 FifteenInc90 TwentyfiveInc90 FiftyInc90 , robust
```

```
inear regression                          Number of obs    =      9,560
                                          F(13, 9546)      =      43.45
                                          Prob > F         =     0.0000
                                          R-squared        =     0.0528
                                          Root MSE         =     4.8258
```

BMI90	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.841018	.2577825	7.14	0.000	1.33571	2.346327
YHI90	.5042238	.1350512	3.73	0.000	.2394948	.7689529
Age90	.0996368	.0217312	4.58	0.000	.057039	.1422346
Hisp	1.323897	.1418888	9.33	0.000	1.045765	1.602029
Black	1.524793	.1269424	12.01	0.000	1.275959	1.773627
Fem	-1.047185	.1051913	-9.96	0.000	-1.253382	-.8409878
OneHS90	-.1706795	.1575434	-1.08	0.279	-.4794979	.138139
TwoHS90	-.3572081	.1787749	-2.00	0.046	-.7076448	-.0067714
ThreeHS90	-1.322378	.1810069	-7.31	0.000	-1.677189	-.9675656
OneInc90	.1388644	.1605862	0.86	0.387	-.1759186	.4536475
FifteenInc90	.1020649	.1671284	0.61	0.541	-.2255423	.4296721
wentyfiveInc90	.2199641	.1702242	1.29	0.196	-.1137115	.5536397
FiftyInc90	.0480345	.2605115	0.18	0.854	-.4626235	.5586924
_cons	22.24872	.6572465	33.85	0.000	20.96038	23.53706

(Input Interaction terms - 1990)
This indicates that with the included controls, females with private health insurance will have a BMI of 0.578 less than a male without health insurance. A female with Medicaid will have a BMI higher by 1.406 than a male without health insurance.

The biggest change is on the females with private health insurance variable. It becomes -0.667 with the inclusion of the interaction term and addition of more independent variables.

(Input Interaction terms - 2010)
Females with Medicaid have a higher BMI of 0.407 in comparison to males with no health insurance, how it is not significant. Females who are private health insurance have a decrease in BMI of -1.739. This is significant.

When more variables are added to this regression that includes the interaction terms, the coefficients on BMI decrease for our selected variables.

Females with Medicaid have BMIs that increase slightly with the inclusion of the interaction term and additional variables. Females with Health Insurance continues to decrease with new variables.

BMI, Medicaid, Age, Gender, Ethnicity, Income, and Education in 2010
Medicaid continues to drop in this regression. Standard error and the F-statistic both decrease because of the addition of new variables. As an individual increase his or her education level, his or her BMI continues to decrease, while being not being statistically significant. Medicaid, Health Insurance, Hispanic, Black, and Some HS are the statistically significant variables in this regression. Income does not have a statistically significant correlation to BMI.

With the addition of new variables, 1,300 observations were dropped. Medicaid correlation to BMI decreases with the addition of new variables, as does Health Insurance correlation to BMI.

More educated individuals are assumed to be more aware of the health-related risks associated with obesity. Education is statistically significant at the 95% level in comparison to the No HS reference group.

```
. regress BMI10 YMed10 YHI10 Age10 Hisp Black Fem OneHS10 TwoHS10 ThreeHS10 OneInc10 FifteenInc10 TwentyfiveInc10 FiftyInc10 , robust

Linear regression                              Number of obs   =      5,588
                                               F(13, 5574)     =      12.79
                                               Prob > F        =     0.0000
                                               R-squared       =     0.0295
                                               Root MSE        =     5.8578
```

BMI10	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.505435	.4907627	3.07	0.002	.5433488	2.467521
YHI10	.467052	.2398745	1.95	0.052	-.0031955	.9372994
Age10	-.0270478	.0343704	-0.79	0.431	-.0944272	.0403317
Hisp	1.395123	.2042545	6.83	0.000	.994705	1.795542
Black	2.020174	.1961233	10.30	0.000	1.635696	2.404652
Fem	-.1593473	.1660743	-0.96	0.337	-.4849176	.1662231
OneHS10	.4990452	.2073975	2.41	0.016	.0924653	.9056251
TwoHS10	.2780074	.3026966	0.92	0.358	-.315396	.8714107
ThreeHS10	-.1051296	.2601269	-0.40	0.686	-.6150798	.4048205
OneInc10	-.1551042	.3980853	-0.39	0.697	-.9355065	.625298
FifteenInc10	.3596232	.3902631	0.92	0.357	-.4054445	1.124691
TwentyfiveInc10	.2704061	.3520494	0.77	0.442	-.4197479	.9605602
FiftyInc10	.2017103	.3652462	0.55	0.581	-.5143146	.9177353
_cons	28.68695	1.710653	16.77	0.000	25.3334	32.04049

BMI, Medicaid, Age, Gender, Ethnicity, Income, Education, Health Insurance, Region, Married, Dependents and Geographics in 1990

In this regression, having Medicaid correlates to an increase in BMI by 1.855. An additional year of age correlates to an increase in BMI by 0.111. Being Female decreases BMI by -0.987 in comparison to the Male baseline. Being Hispanic correlates to an increase in BMI by 1.541 and being Black correlates to an increase in BMI by 1.601, in comparison to the White baseline. In this regression, there is no correlation between BMI and Income, it seems to be scattered here, due to the Edu effect. Income and Education are related to each other. If an individual has a higher income, they have access to a better education and vice versa.

Being married correlates to an increase in an individual's BMI by 0.707, however, not being married correlates to an increase in an individual's BMI by 0.666. Living in a rural environment correlates to an increase in BMI by 0.508 in comparison to the urban baseline. It can be assumed that individuals who live in a city walk more because everything is within walking distance. In terms of the number of dependents one has, there is no statistical significance.

Geographics does play a role in one's BMI, with the Northeast region having the highest BMI of a coefficient of 0.407, with the West as the baseline. Region shows a positive correlation in BMI, with the other regions having a coefficient of 0.336 and 0.287 for North Central and South, respectively. Jeffrey Kling and Jeffrey Liebman's research, illustrates why region plays a large role in BMI. Their study looks at the causal effects of teenage youth moving out of high poverty neighborhoods. Youth who grow up in disadvantaged neighborhoods fare poorer health outcomes than those who grow up in affluent neighborhoods.

```
regress BMI90 YMed90 YHI90 Age90 Hisp Black Fem OneHS90 TwoHS90 ThreeHS90 OneInc90 FifteenInc90 TwentyfiveInc90 FiftyInc90 Rur90 YMarr90 NMarr90 ZDep90 OneDep90 TwoDep90 NEast!
0 NCen90 Sou90 OneOcc90 TwoOcc90 ThreeOcc90 , robust
```

linear regression				Number of obs	=	9,246
				F(25, 9220)	=	23.88
				Prob > F	=	0.0000
				R-squared	=	0.0576
				Root MSE	=	4.8535

BMI90	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed90	1.825463	.2670686	6.84	0.000	1.301949	2.348976
YHI90	.4772537	.1434227	3.33	0.001	.1961134	.758394
Age90	.1121596	.02369	4.73	0.000	.0657219	.1585974
Hisp	1.54427	.1551368	9.95	0.000	1.240167	1.848372
Black	1.602699	.1388292	11.54	0.000	1.330563	1.874835
Fem	-.990081	.1125479	-8.80	0.000	-1.2107	-.7694622
OneHS90	-.1638769	.1610698	-1.02	0.309	-.4796094	.1518556
TwoHS90	-.3307891	.187135	-1.77	0.077	-.6976152	.0360369
ThreeHS90	-1.384104	.1956852	-7.07	0.000	-1.76769	-1.000517
OneInc90	.1677398	.2390783	0.70	0.483	-.3009066	.6363862
FifteenInc90	.1281941	.2442789	0.52	0.600	-.3506466	.6070348
TwentyfiveInc90	.306961	.2466718	1.24	0.213	-.1765704	.7904924
FiftyInc90	.1524624	.3169419	0.48	0.630	-.4688137	.7737386
Rur90	.5082653	.1346204	3.78	0.000	.2443794	.7721512
YMarr90	.7063098	.1494072	4.73	0.000	.4134386	.9991809
NMarr90	.6623627	.1709179	3.88	0.000	.3273257	.9973996
ZDep90	.0508009	.190439	0.27	0.790	-.3225017	.4241035
OneDep90	.0168089	.1786084	0.09	0.925	-.3333032	.3669207
TwoDep90	-.180416	.1729948	-1.04	0.297	-.5195241	.1586922
NEast90	.4052972	.1640818	2.47	0.014	.0836605	.7269339
NCen90	.3360665	.1591676	2.11	0.035	.0240628	.6480703
Sou90	.2859533	.1466079	1.95	0.051	-.0014306	.5733372
OneOcc90	-.2894292	.2799092	-1.03	0.301	-.8381132	.2592549
TwoOcc90	-.1662191	.2966002	-0.56	0.575	-.7476211	.4151829
ThreeOcc90	-.1858274	.3313218	-0.56	0.575	-.8352914	.4636366
_cons	21.10417	.777178	27.15	0.000	19.58073	22.62761

BMI, Medicaid, Age, Gender, Ethnicity, Income, Education, Health Insurance, Region, Married, and Geographics in 2010

Having Medicaid correlates to an increase in BMI by 1.606. Being Hispanic correlates to an increase in BMI by 1.566 and being Black correlates to an increase in 1.948, in comparison to White individuals. One's BMI decreases with increased education, however, only Some HS is statistically significant, with Some HS correlating to an increase in BMI by 0.512.

Marriage positively affects BMI. Not being married increases BMI the most, comparatively. With the Region variable, only South was statistically significant, correlating to an increase in BMI by 0.608. Only having one dependent correlates to a decrease in BMI by 0.531.

```
regress BMI10 YMed10 YHI10 Age10 Hisp Black Fem OneHS10 TwoHS10 ThreeHS10 OneInc10 FifteenInc10 TwentyfiveInc10 FiftyInc10 Rur10 YMarr10 NMarr10 ZDep10 OneDep10 TwoDep10 NEast1
0 NCen10 Sou10 OneOcc10 TwoOcc10 ThreeOcc10 , robust
```

linear regression				Number of obs	=	5,552
				F(25, 5526)	=	7.88
				Prob > F	=	0.0000
				R-squared	=	0.0354
				Root MSE	=	5.8509

BMI10	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
YMed10	1.485674	.4949105	3.00	0.003	.5154540	2.455893
YHI10	.451404	.2471466	1.83	0.068	-.0331004	.9359085
Age10	-.0265688	.0345613	-0.77	0.442	-.0943235	.041185
Hisp	1.567286	.2273101	6.89	0.000	1.121669	2.012903
Black	1.950179	.2168893	8.99	0.000	1.524991	2.375368
Fem	-.1128316	.1675065	-0.67	0.501	-.4412101	.215547
OneHS10	.5294422	.2071991	2.56	0.011	.1232505	.9356339
TwoHS10	.2959764	.3004531	0.97	0.333	-.3028319	.8947847
ThreeHS10	-.1195345	.2618449	-0.46	0.648	-.6328535	.3937845
OneInc10	.1443176	.4413625	0.33	0.744	-.7207305	1.009366
FifteenInc10	.6599195	.4415613	1.49	0.135	-.2057143	1.525553
TwentyfiveInc10	.6061078	.4089709	1.48	0.138	-.195636	1.407052
FiftyInc10	.5739083	.4190696	1.37	0.171	-.247633	1.39545
Rur10	.379137	.2048394	1.85	0.064	-.0224287	.7807027
YMarr10	.462212	.1926736	2.40	0.016	.0844959	.8399282
NMarr10	.6903905	.2930779	2.36	0.019	.1158426	1.264938
ZDep10	.1600138	.2565721	0.62	0.533	-.3429683	.662996
OneDep10	-.5335672	.2401201	-2.22	0.026	-1.004297	-.0628373
TwoDep10	-.2017762	.1945142	-1.04	0.300	-.5831005	.1795481
NEast10	.1601892	.2669333	0.60	0.548	-.363105	.6834834
NCen10	.3301524	.25192	1.31	0.190	-.1637098	.8240146
Sou10	.6085766	.227177	2.68	0.007	.1632203	1.053933
OneOcc10	-.5932016	.5058651	-1.17	0.241	-1.584896	.3904929
TwoOcc10	-.3649223	.5347566	-0.68	0.495	-1.413256	.683411
ThreeOcc10	-.4484688	.5854847	-0.77	0.444	-1.596249	.6993116
_cons	28.14955	1.756111	16.03	0.000	24.70688	31.59222

BMI, Medicaid, Age, Gender, Ethnicity, Income, Education, Health Insurance, Region, Married, Geographics, and Occupation in 1990

With the addition of the occupation variable in 1990, none of the other variables are statistically significant.

**BMI, Medicaid, Age, Gender, Ethnicity, Income, Education, Health Insurance, Region, Married, Geographics, and Occupation in 2010**

With the addition of the occupation variable in 2010, none of the other variables are statistically significant.

BMI and Income in 1990

A scatter plot of BMI versus Income in 1990 indicates some outliers, or some top-coded income. A trend in 1990 portray that lower income is positively correlated to lower BMI. Income is maximized at $75,000. BMI above 50 is highly unlikely (since average BMI is 26.6 for adult males and 26.5 for adult females, and anything above a BMI of 30 is obese, according to the CDC). After analyzing the scatter plot, BMI above 50 will most likely be dropped.
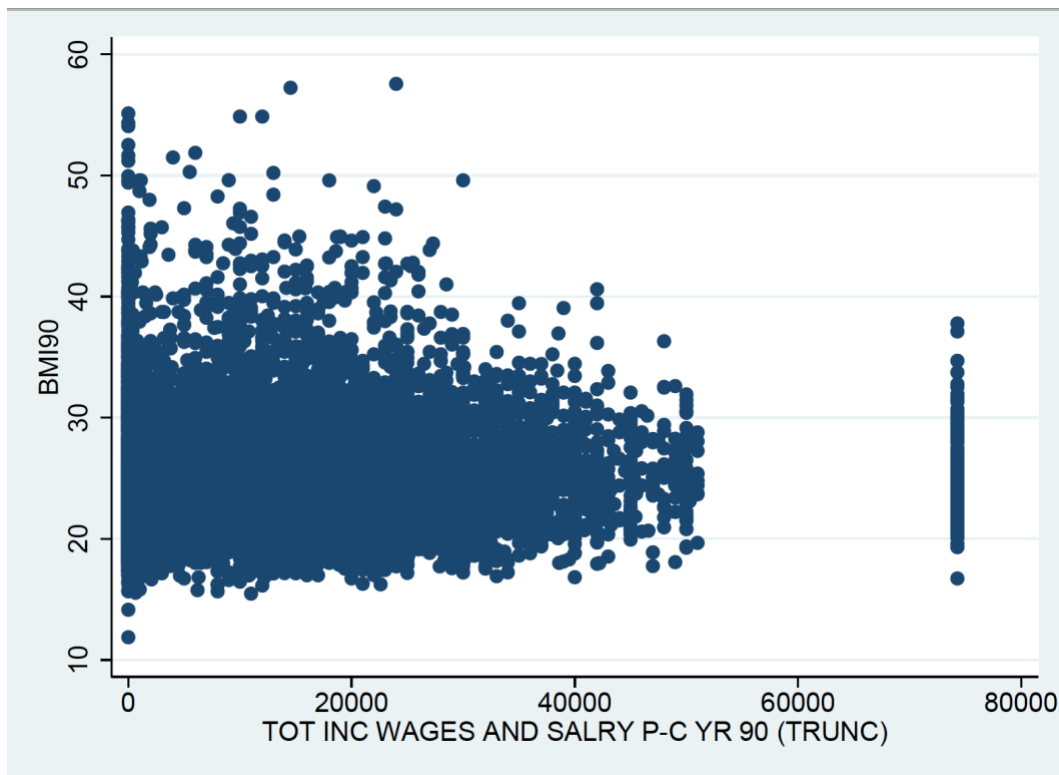
BMI and Income in 2010

In the scatter plot of BMI and Income in 2010, Income is maximized at $300,000. Most individuals who have an income of $300,000 or more do not fall into the overweight category. The trend of BMI and Income in 2010 indicate that as Income increases BMI will decrease. There are no data points between the Income range of $150,000 and $300,000. This could be due to top incomes given the average in the group.

Contrasting BMI and Income in 1990 versus in 2010 indicates that there is a difference in Income caps, with $75,000 for 1990 and $312,000 in 2010. The significant difference in maximum Income levels can be attributed to inflation and cost of living adjustments over the 20 year time period.
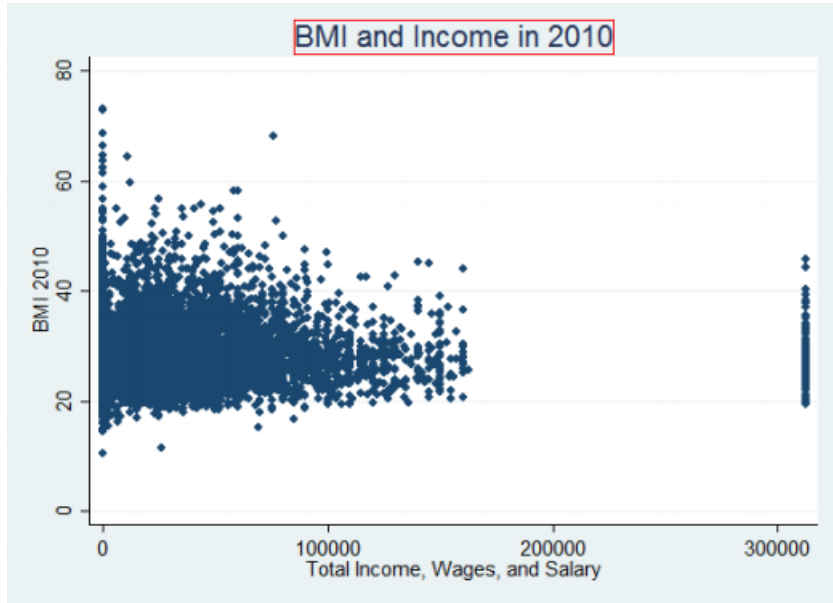
Bysort Obese and BMI in 1990 and 2010
When Obese and BMI are subgrouped by their respective years, it can be noted that a greater percentage of people surveyed in 2010 were more obese in 1990. In 1990, 14.8% (1,488/10,037) of the surveyed population were obese with an average BMI of 34.31 and a max of 57.56. In 2010, the percentage of obese people doubled in 20 years to 37.14% (2,721/7,237). While the mean increased slightly to 35.30, the max increased to 73.15. While the maximum is only one person, the drastic increase in maximums in 20 years could point the continued trend of increased BMI.