Does Obesity Really Reduce Wage?

MEMORANDUM

TO:

United States DEPARTMENT OF LABOR

From:
Tomer Eldor[1]
Minerva Schools
April 2017

1 Corresponding Author: For any comments, questions, getting data and code for replication, contact the author via to@minerva.kgi.edu.

EXECUTIVE SUMMARY

Recent calls from the Obesity Action Coallition and Obeisty Society claimed that obese employees receive lower wages than their counterparts. This is theorized to happen because obesity increases health insurance costs for the employer, and these costs then fall on the worker as lower wages. The ministry of labor needs to know the validity of these claims to decide of designing policy.

I. A study by Jay Bhattacharya and Kate Bundorf (2010) suggested that obese employees with employer-sponsored health insurance recieve lower cash wages. They find the penalization bigger for women, and biggest for white women. They also compared workers without insurance and find no observable difference between obese workers and not. However, this study analyzed the groups as they were without balancing participants' attributes across groups. That means that their observable differences might be due to inherent differences of people between groups. As obesity is associated with lower socioeconomic class, their wages might be lower because of other factors like lower education level or racial discrimination; thus their detected effects might be invalid.

II. Anas El Turabi and Phil Saynisch (2014) replicate and extend their findings using Mahalanobis Distance Matching – a method aiming to balance the groups in terms of their underlying attributes, and by examining the effects for subgroups by race-gender pairs. They supported the original findings, and their partitioning revealed that the largest effect was 11$ penalization for white women. However, their Matching technique is not optimal, so their data is still not greatly balanced and their results might be confounded too.

This paper replicates the methods and results of T&S and extends them with the following:

Analysis Conducted	Conclusion
1. Achieveing more valid and credible estimated effects, by performing Genetic Matching on each gender-race group, seperately. This greatly improved their balance, and yielded more credible results.	→ **Obesity does not reduce wage**. We found no (significant) effect for any race-gender pair group. This contradicts previous findings of penalization, suggesting these where observed merely due to other underlying unbalanced group differences.
2. Evaluating differences associated with having employer insurance on wages, seperately for obese and non-obese workers.	→ **Insured obese and non-obese workers earned similarly more** ($4.5 and $4.2, respectively) than their non-insured counterparts.
3. Inspecting trends of wage associated with continuous BMI at different quantiles of wages.	→ **Women wages slightly decreased with BMI**. The higher they earning quantile was, the sharper the decrease; While **Men's wages marginally increased** with BMI. However, the effects were still small.
4. Analsying discontinuity in wages for white women around the "Obesity" threshold, BMI of 30.	→ There was an **average drop of 2$ around BMI of 30** but it was still **statistically insignificant**.

Reccomendations. This more rigorous analysis reveals that there is no really penalization of wages for insured obese workers. The ministry should not design policy to counteract this phenomena yet, but rather meet with stakeholders and refine their claims, and potentially get a deeper look.

BACKGROUND INFORMATION: FOUND IN APPENDIX

WHAT DID THE PREVIOUS PAPERS DO, AND WHY IS IT PROBLEMATIC

B&B use the NLSY data to estimate the impact of obesity on healthcare costs, using difference-in-differences technique. They conclude that healthcare obesity costs *are transferred to obese employees* by a wage reduction of $1.45 per hour, after controlling for employee characteristics. However, they did not match covariates properly. My simple balance check of the untreated data showed that it is highly unbalanced, with most covariates having p-values below 1%, with 4 variables having the zero-approaching minimal blaance : a Before Matching Minimum p.value smaller than < 2.22e-16, where some of these variables, i.e. education and financial industry, are correlated with wages, so having them unbalanced within groups induces bias in wage estimation – mostly overestimation because of key imbalanced covariates which are acossiated with the outcome. Therefore, I will perform matching of these groups and redo the difference of differences to get a less biased result.

T&S built on B&B's results, performed Mahalanobis-Distance Mathcing on the dataset, divided the data into 4 subgroups according to gender and race, estimated the local effect and other statistics. I hereby replicated their results. Fortunately, their data and code were easily available and replicable.

Replication of T&S

Replicating T&S process yielded very similar results but not perfectly the same. When replicating, MD matching improved balance for only ***20 covariates,*** versus 23 covariates in T&S. This means that the groups might have been even less balanced than reported, and lowers the internal validity of T&S. The resulting outcomes are also slightly different, perhaps because of the difference in matching balance, but not in any way changing their conclusions. See Appendix for full replication.

Limitations of T&S study

The study by T&S, while improving the rigor of the study by B&B, did have limitations.

First, Mahalanobis-Distance Mathcing is a far from ideal way to match. Mahalanobis distance is a scalar quantity measuring the Euclidean distance[2] and scaling it by the standard deviation. However, it is limited for multiple reasons:

---

[2] Mahalanobis Distance Measure Distance between observations basing on covaraites; between Treatment unit Xt and the nearest control unit :MD(Xc,Xt) = √(Xc-Xt)'S-1(Xc-Xt).
Adapted from:
Diamond, Alexis & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, *95*(3), 932-945. Retrieved from https://service.sipx.com/service/php/inspect_document.php?id=perma-x-c66c955a-5f28-11e6-a73e-22000b61898b

1. It does not minimize the loss function
2. It treats all values with equal weight (given variance)
3. It generates a wrong scaling of each variable. It is highly affected by the original scale and unit of the vector. For example, if we measured age by milliseconds rather than years, it would skew our results to match on increasingly close age, trading-off the importance of other covariates.
4. As stated by Diamond and Sekhon (2012), Mahalanobis Distance does not perform well when covariates have non-ellipsoidal distributions,
5. By using MD, as is with propensity scores, we lose valuable information about the individual distance from each covariate.

Therefore, their matching is not ideal. It is crucial to have good balance here, especially since many covariates are correlated with wage outcomes, beyond race and gender: education level, industry, occupation, having children and living in a city. Their p-values after MD matching in T&S were mostly still low – all below 0.4 and some too small to specify (<0.001), meaning that groups were unbalanced on covariates correlated with outcomes, so that there will likely be inherent differences in wages solely for these underlying covariates, and not the treatment. For example, the non-obese group may have higher-educated, city residents with higher paying occupations and industries, and thus we will observe differences between the obesee and non-obese groups – but that is not the effect of obesity. However, there is a solution: Genetic Matching. GenMatch() fucntion, developed by Diamond and Sekhon, uses a genetic algorithm to obtaim optimal weights for matching and improves balance greatly. Therefore, I used genetic matching for similar inspection of subgroups to see if T&S reported effects still hold with a better covariate balance.

T&S produce clear and coherent tables comparing effect estimates of obesity on wages within differennt groups. However, to do so, they use *simple* linear regression on the unmatched dataset. You may see that in the code replication section, by inspecting each "model", i.e. "modelwmale", observing that it is simplly a linear model (*lm()*) on the *unmatched* dataset: "nlsy" instead of the matched "data.mahal". Therefore, these estimates are likely biased by covariate imbalance acossiated with outcomes (i.e., inherent differneces in education, location, etc.), rather than of obesity, thus overestimating the effect.

Data & Methods

I used the The National Longitudinal Survey of Youth (NLSY), used by B&B and T&S. This dataset is vast, with above 31,000 observations of 12,686 invidividuals, analyzing 35 of their recorded variables.[3] It did not come without limiations, since as a survey, the measurements our self-reported,

[3] NLSY is a nationally representative longitudinal survey in U.S.A based on a cohort of 12,686 individuals that were aged between 14-22 in 1979. It collects dozens of variables; B&B, T&S and I here use approximately up to 35 of them related to weight, income, and covariates related with these two.

and thus might be incorrectly reported, intentionally or due to measurement error; and it involved list-wise deletions of measurments. Therefore, we have irreducible error leading to uncertainty.[4]

GENETIC MATCHING FOR ESTIMATING THE TRUE EFFECT *SEE APPENDIX FOR A BROADER DISCUSSION

To improve the internal and inferential validity of the estimated effects of interest, I applied genetic matching instead of Mahalanobis Distance matching in T&S (and instead of no-matching at all in B&B). First, before that, I ran linear regression on obesity and then on wage onto all other variables, In order to: (1) inspect the validity and relevance for subdividing the data according to race in gender such as was done in T&S without rigorous justification; and (2) determine which covariates would be possible to not match on, because their unusually large number defeats our ability to match well on all covariates. I found race and gender to be amongst the highest correlated, and thus continued to examine like T&S in race-gender pairs. There was one difference: T&S dropped "other" race observations. This limits the external generalizability and analysis of the study. I found that it is as correlated with obesity and wages as "black" is. I therefore included it with the "black" women group (where there was even *higher correlation)*. However, since they claimed for a significant effect for black men later, I compared only black men later too.
Another difference from T&S study was *when* we matched. T&S matched the entire sample first, then divided it into groups; which reduces the balance more.[5] I *first divided the data into gender-race pairs, then* applied genetic matching for obese versus non-obese groups within each strata.
A final difference was that I generated propensity scores for each individual and matched them too.

BALANCE IMPROVEMENTS DUE TO GENETIC MATCHING

The balance of covariates (attributes) improved significantly using genetic matching, as seen in Table 1 below. This was achieved due to the algorithmic advantages of genetic matching, and also by subdividing to groups first and matching each subgroup individually to maximize its otpimality. This more balanced data yields more credible, internally valid, and less biased results.
The analyzed groups here were only for people with insurance, as were the relevant estimates discussed in T&S, since the question is irrelevant for people without insurance, and it would have confounded our results.

[4] A bit about uncertainty. The true effects are hard to tell. Without direct counterfactuals of how would this person earn in the same conditions if they were'nt obese, can never know the truth. Complex personal and social factors drive personal and market outcomes which even when observing, we are not necessarily aware of. We always have some irreducible error we can't know. Apart from the mentioned reporting error, We have exogeneity concerns for hidden variables we don't have data on, but which might have affected both obesity and wage or socioeconomic factor – like peronal motivation, home and environment culture; so if those take place and increase the association of treatment with outcome, the effect will be overestimated.Therefore, we have uncertainty, and I will try to estimate the uncertainty of each estimand I give using confidence intervals for example.
[5] The groups were not well balanced on race and gender before division: Post-Macthing Black covariate was 0.15 for non-obese and 0.17 for obese, with the matching p-value under <0.001. This means that after dividing the data, the respective obese vs. non-obese groups are even less balanced then before.

Covariate Balance: (Minimum P.Value)	No Matching	T&S: MD Matching	New: Genetic Matching
Black Women / NonWhite Women[6]	P < 2.22e-16 (Lowst Possible) 3 vars: age, srvy_yr, educ (+propensities)	P: <0.001 3 vars: Urban Residence, tenure, survey year. Importantly, the balance of uninsured and epmployer insurance was 2nd lowest: P: 0.002	P: 0.001 1 var: Childany
White Women	P < 2.22e-16 (Lowst Possible) 3 vars: age, srvy_yr, educ (+propensities)	P: <0.001 3 vars: Urban Residence, tenure, survey year.	P: 0.0759 1 var: Childany
Black Men	P < 2.22e-16 (Lowst Possible) 5 vars: childany, age, srvy_yr, educ (+propensities)	P: <0.001 3 vars: Urban Residence, tenure, survey year.	P: 0.024569 1 var: childany
White Men	P < 2.22e-16 (Lowst Possible) 3 vars: age, srvy_yr, educ (+propensities)	P: <0.001 3 vars: Urban Residence, tenure, survey year.	P: 0.005667 1 var: childany

TALBE 1: Covariates Balance Comparison

The results showed clear improvement in balance in covariates. However, we shuold note that in our matched sample, one repeating variable wasn't matched well: ChildAny – if the subject has any children. This means out resutls might be confounded by this imbalance and shuold be noticed. However, this variable should not be a direct determinant of wage, and was problematic to match on in the firstplace, since we could claim that it is a consequence of obeisty: due to social appearance-based discrimination, obese people might be less likely to have children because of their obesity; but this is not proven, therefore it was better to try to match on this variable, but its imbalance is still better than previous balances. We can proceed to show estimated effects of higher validity.

COMPARING ESTIMATED EFFECTS OF OBESITY ON WAGE

After dividing for race-gender pairs, within the people insured by employers, performing Genetic Matching to balance covariates, we can see the estimated "effects" of obesity and their uncertainty. The estimated effects clearly tell a different story than the original; there is no difference between wages of obese and non obese people with employee insurance – for niether race-gender based subgroup, including white women. This suggests that the observed effects in previous studies are

[6] Black Women / NonWhite Women are *similar* but *not Identical* groups. The total sample of Nonwhite women, before matching, was composed out of the 3488 black women included in the study, and 594 women of "other" races (other minorities than black, which correlated similarly with obesity and wages as black women did). See full explanation under "Examining Correlation of Obeisty and Wage to Other Variables" section.

merely due to confounding differences between the groups in their analysis and *not* due to obesity, since these other confounding variables weren't controlled for as well in previous studies. Sensitivity analysis were also run, and the first p-value associated with an increase of change (Gamma) of 0.1 are reported in the table; yet since the estimates were already insignificant, the p-values associated were already beyoned the 5% significant rate, so they bear no meaningful results

EFFECT ESTIMATES (in $ per hour)	**GenMatch Penalty**	SE	Confi-dence Interval	Conf-idence Interval	T&S Estimate	T&S SE	T&S Signi-ficant	GenMatch: Significant	Sensitivity Analysis: Estimate Gamma1
Black Women / NonWhite	**-0.188**	0.292	-1.29	+0.37	-0.742	0.542	No	**No**	0.998
White Women	**0.551**	0.364	-0.162	1.265	-3.104**	1.360	Yes	**No**	0.9987
Black Men	**-0.138**	0.352	-0.829	0.553	-0.474	0.673	No	**No**	0.9411
White Men	**0.005**	0.240	-0.087	0.729	-0.562	0.779	No	**No**	0.2203

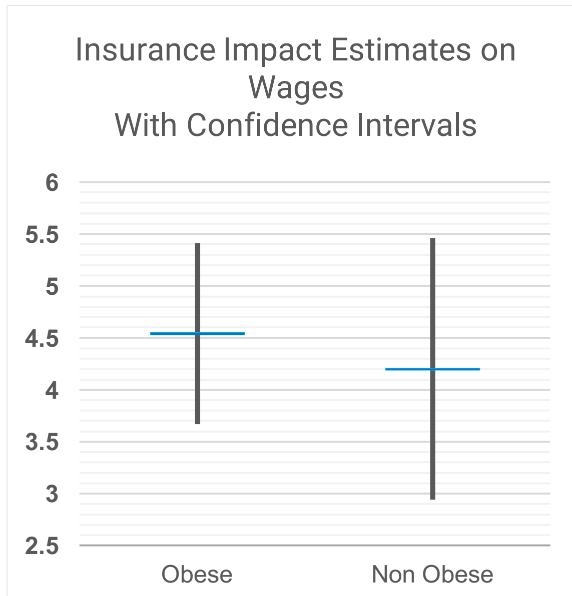
These results suggest that we can contradict prior concerns, and we should not think that obesity leads to lower wages in employee-insured workers, nor for white women or any other group. According to this dataset, you can dismiss concerns about penalizing obese employees; or else, start being concerned about the potential moral hazard promoting obesity related.

Does Having Employee Insurance Impact Wage?

The underlying hypothesis in papers like B&B and T&S it that *due to the increased insurance costs for the company, it pays obese employees less.* T&S do not offer a rigorous inspection or discussion about whether the wage penalization is related to having employer insurance, or is the wage lower for all obese workers because of pure discrimination? B&B did inspect that, but they did not match covarites, so their estimate might be biased.
This effect is even more relevant than the effect of obesity; since obesity is not something easy to change, and more like a trait; but employee-covered insurance could be considered a proper

treatement which can be manipulated. Hence it is more appropriate to speak about that as treatment.
I divided the data into obese and non obese groups. Within each of those, I examined the difference between the subgroups: with insurance, or without insurance – by performing **Genetic Matching**.
The balance increased greatly,
Examining the obese group shows how insurance affected wages for obese people. Yet there might be differences in wages related with insurance and unrelated to obesity, such as insurance included jobs tend to be by wealthier, bigger companies, with higher wages. Therefore, I examined also the baseline difference in wages by insurance within the non-obese group, after genetic matching.
The results show that **both obese and non-obese groups earned similarly more, in average, *with the insurance* than without:** $4.54 and $4.20 more per hour, respectively, with highly overlapping confidence intervals.

Impact of Insurance in $	Obese	NonObese:
Estimate	4.54	4.20
Confidence Intervals of 95%:		
Upper Bound	3.66	2.94
Lower Bound	5.41	5.46



Sensitivity analyses, (at end), were run, and showed the estimates are significant and relatively insensitive to exogeneity (hidden bias or unobserved variables). Yet since the difference between the groups are not statistically significant, sensitivity tests have no meaningful results in terms of the difference. The resutling confidence intervals get wider respectivly and overlap even more.

Conclusion: Insurance Does Not Impact Obese Employees' Wage More Than Usual

The result is that insurance seems to be associated with increased wages, but the estimated effect is not different for obese people than it is for non-obese people. Hence, there is no different treatment of obese employees than there is usually. Why is there increase in average pay for workers with insurance? I suggest they might be simply bigger, wealtheir compnies with higher wages normally, as said previously.

A Closer Look: How are Wages Distributed by BMI?

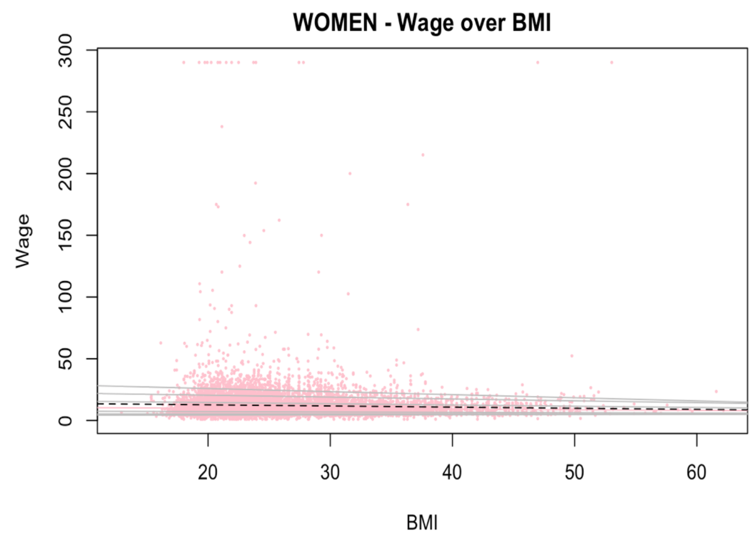
So far, we have only looked on binary groups of “obese” versus “non-obese”. To inspect if there is any difference as the BMI changes, and if the trends of wage is difference for people with lower wages than for people with higher wages, we turn to look at continuous regressions.

ARE THE TRENDS AMONGST LOWER WAGES DIFFERENCE THAN IN HIGHER WAGES?

To examine such wage-quantile differences, we run Quantile Regression.[7] As the data was divideed by gender before, we do so here, running seperately for women and men.

Quantile Regression for Women

This Regression reveals a few things: first, most wages are concnetrated between 0 and ~30, with few outliers of high wages that obscure the patterns. Therefore, considering quantiles and medians, or even referring just to the effects for the rest 90% is beneficial, and we will do that next. Second, the median effect line has a marginally decreasing trend.
The lowest quantiles have almost no effect at all. This is because there is a natural limit to how low wages go; and there are people with similarly lowest wages at any weight. As we the quantiles get higher, the effect is a greater decrease. We can see that more clearly, with less outlier interference, on the right with plotting the *Log* of wage over BMI. However, this trend is still small, still not significant, hence not enough to act on.

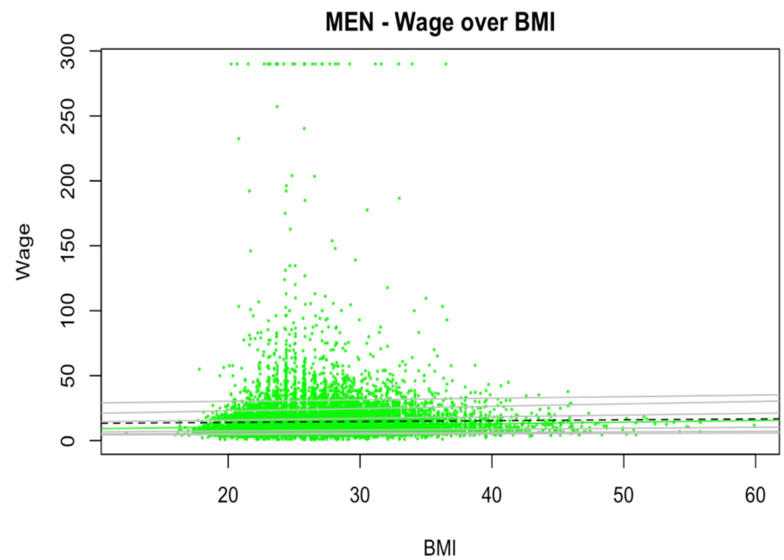


Quantile Effects - the quantile of differences between obese and non-obese groups were examined (detailed at end). The most interesting result was for white women, where all 5% quantiles, presented

[7] Quantile regression allows us to see how different *types* of participants were impacted; in this case, how was did BMI (indicator of body fat and thus obesity) influencing people of lower to higher wages; at the different quantiles of the distribution of outcomes, providing a more complete picture of their relationship. It is also more robust to outcome outliers.

at the Appendix, were slighlty negative, between -$0.8 and -$2, except for the lowest and highest quantiles containing the outliers, which had high positive values, but aren't representative. However, these reductions are small and not necessarily significant. Therefore, if the specific quantiles of white women are of interest for you, further examination is needed to understand the exact effect, confidence and significance with more robustness. However, since you are interested at policy at large, quantiles of one quartile are not representative, and I would not design wide-ranging policies from that.

Quantile Regression For Men



Examining the effects for men reveals firstly interesting results: the trendlines show a slight *increase!* But this increase is also still very small and thus not significant. Nonetheless, it is interesting to note the opposite directions in trends for genders. As they both are small, it might be simply correlation by chance and not causation. However, one may develop hypotheses to explain such phenomena: for example, there exists a stereotype of men who are wealthy and well-overweight, particularily at high-paying jobs such as the financial-sector, and it has somewhat become more socially acceptable. Converstly, this stereotype is not common for women, and they might be more adversly hurt by discrimination based on weight on top of gender discrimination, especially by male managers. This is only a first line of hypothesis for investigation, and if it is of interest, this should call for continued research.

IS THE CUTOFF OF "OBESITY" AT BMI=30 REAL?

Our last refinement concerns with the continuity of BMI versus the binary definition of "Obese". Obesity is defined as having BMI of above 30; yet in reality, weight and BMI are continuous scales, and we can not detect visually any harsh cutoff. However, if health insurance companies do measure the BMI of employees before suggesting premiums, then this "cutoff" bares weight. We will examine the trend around this "cutoff" using Regression discontuity analysis. We will focus only on white women this time, since that was the only originally potentially significant difference and topic at question. If we find no effect for them, then clearly all others will be even less significant.

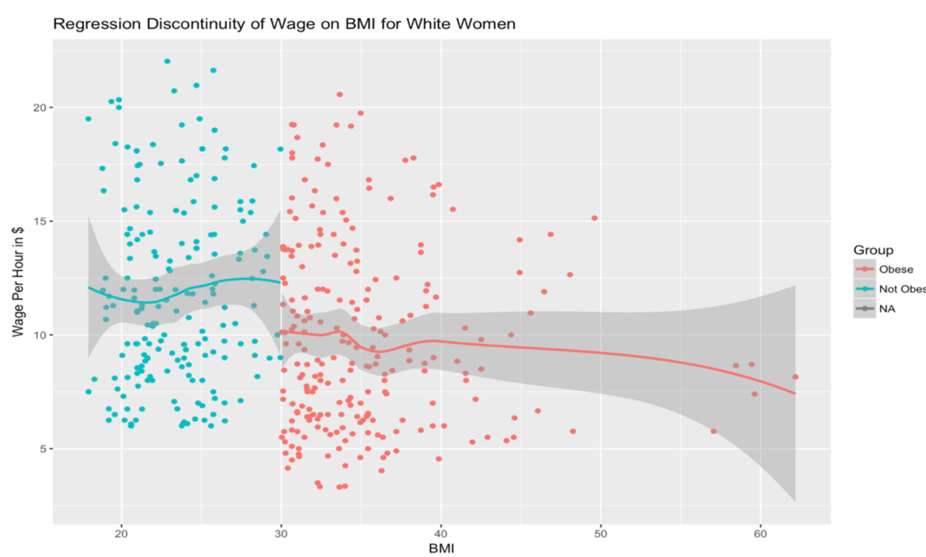
Local Regression and Regression Disconuity for White Women

First, we took the **matched dataset** of obese white women to non-obese white women counterparts, to have an appropriate comparison.

Second, as we saw in previous figures, there were many outliers which skew the trend. Here we are concerned about the pattern for most of the data. These outliers were few and highly skeweing the curves into reflecting inaccurately the bulk 90% of the data. We prefer to accurately reflect 90% of the data. Hence we dropped the outliers with the largest wage per hour - the top 5% quantile, and the 2 observations with unusually small wage (at the bottom 5% quantile).

First, see the figure at right, with continuous local regression (loess) on the two groups. Even here, we see that there ***is*** a sharp decrease arriving to 30; so we examine it more closely with regression disconuity.

Regression Discontinuity shows that while there seems to be an estimated wage decrease of -$2.21 around the "cutoff" of BMI=30, this effect is not significant.



The confidence interval around that effect is between -5.82 and 1.40, therefore we can't say at 95% confidence that if it is even positive or negative, hence it is not significant. We can see that in the plot: although there is a visible cutoff, the confidence intervals around it from both sides overlap; therefore it is not necessarily significant.

Why wouldn't there be an effect?

While people are defined "obese" when their BMI is above 30, in reality there is no visible cuttoff to detect BMI of 29 or 31. Beyond the "definition", there was no other treatment. Regression discontuity designs are best utilized when the "treatment" involved an actual intervention; but in our case, we want to consider the impact of obesity, considering obesity as the treatment – when it isn't an intervention. Therefore, except for BMI, there was nothing really seperating the people just below and just above BMI of 30.

Then how come we see a jump? The confusing pattern might be due to the sensitivity of the manipulated data. I use the Matched data from groups of obese vs non-obese. The matched sample shows a peculiarity: there is less data for BMI between ~25-30, and in particular – there seem to be missing observations of low wages as there are before and after. Exactly there we see a large increase, suggesting that this increase might be increasing for its sensitivity to the fewer, skewed data, and potentially the missing data might have completed the pattern of lower wages and decrease the slope in a more continuous manner. Why is there data missing? This might be because the matching procedure didn't find these units there as good fits, thus dropping more units there. We see that there are far less units at the bottom (with low salaries) just below BMI=30 as there are for all other BMIs (until 40 at least). Thus, the local regression at that point is more sensitive to the observations left, and the lack of bottom units skewed the local regression upwards. Logically, I can think of no good reason for why there would be actually less comparable units at that range with low wages just below 30 than there are in any other BMI range; Nor there seems to be any good reason for people with BMI of just above 30 to be discriminated than from people just below 30. The only potential reason would have been only if they were actually penalized by having employer health insurance for being defined as obese by the insurance; but this would both require the assumptions that employers know their employees exact BMI /obesity definition / individual insurance coverege, which is unlikely, and penalize accordingly; moreover, our examination of the impact of insurance shows the opposite trend! insurance had a positive effect on people's wages.

Conclusions and Next Steps

All analyses show that there is no significant difference in wage based on obesity, for neither race-gender pair group, and that insurance does not related with decreased wages for obese people relative to their non-obese counterparts – but it correlates with higher wages as much as it is for non-obese people. Therefore, obesity, and employer-insurance for obese workers, does not penalize wages when workers are compared to their appropriate counterparts. The differences from previous researches has probably stemmed from imbalance in attributes of their obese and non-obese groups, as obsity is correlated with loer socio-economic class in America, lower education, rural residence, and other confounders, one might examine these as causes.

My recommendation following these analyses is not enacting any policy to mitigate obesity-based wage discrimination, but rather to explain (or send) the findings of this paper to the obesity organizations, and start together a collaborative discussion to see where do they see these issues arise and arrive at *why* together. It may be happenning within certain companies, which would requier analyzing a different dataset including company identifiers, it may be due to individual differences, or it may be just overheared rumors from previous studies with questionable results.
It was a good idea to examine the literature more closely, as you potentially just saved yourselves the funding and work on an meaningless campaign.

References

Bhattacharya, J., & Bundorf, M. K. (2009). The incidence of the healthcare costs of obesity. *Journal of Health Economics*, 28(3), 649–58.

Diamond, Alexis & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics,* 95(3), 932-945. Retrieved from https://service.sipx.com/service/php/inspect_document.php?id=perma-x-c66c955a-5f28-11e6-a73e-22000b61898b

Anas El Turabi and Phil Saynisch, 2014, Overweight and Overburdened: Race and Gender Disparities in the Incidence of the Healthcare Costs of Obesity, with Replication data, doi:10.7910/DVN/25680*, Harvard Dataverse, V2,* UNF:5:0essvnN8MG7uMVfURDp0lw==

NLS. (2017). NLSY Dataset Variables Investigaor. *Nlsinfo.org* Retrieved 21 April 2017, from https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY97 - *used to investigate and understand varaibles in the NLSY dataset.*

Appendix 1: Further Background Information

WHY ARE OBESITY COSTS OF INTEREST

Being obese doesn't only cost that individual with own physical health (and often, subjective wellbeing, due to societal response) – it also increases healthcare costs for that individual. In the aggregate, in a U.S population that is getting increasingly obese, these costs would only increase. Nowadays, healthcare insurance costs is often funded not by the individual, but through employer insurance (using funds which might have been allocated to better serve other customers) or governmental subsidies (such as Obama Care). Therefore, obesity increases the healthcare costs for insurance companies, but who pays for that cost burdon? Is it the firms themselves? the obese workers? All the employees? Or the insurance company's customers, like you or any reader? This question is of great relevance for healthcare policy making, insurance companies, firms, worker rights unions, and anti-discrimination interest groups. Moreover, since these costs are at direct tradeoff with other potentially more helpful expenditure for you, this issue might be directly relevant to any reader.[8]

BACKGROUND: THE THINKING SO FAR: ORIGINAL PAPERS

The underlying theory in T&S and B&B[9] is that penalization of wage because of obesity are due primarily to higher costs for the employer-sponsored insurance. Having more obese workers would cost the firm more, but obese employees enjoy free insurance, when they need it more. T&S raise the concern that this might create moral hazard – promoting obesity without baring the healthcare costs yourself. As T&S demonstrate, lowering wages for obese employees is even more adverse for the created reinforcing feedback loop between poverty and obesity; obesity is much more common in poorer status, and since fattenning, unhealthy food is cheaper in the U.S, Drewnowski & Specter (2004) suggested poverty *causes* obesity. Therefore, if wage decrease falls on groups *already at higher risk of obesity* – which is highly correlated with groups correlated with lower-wages, this would create a reincoring feedback loop of lowered wages and increased weight. Moreover, if reduced wage seemed as a just "penalty" for the costlier obese employees, would just be getting them deeper in that mud, promoting increased obesity.

---

[8] If you are interested in understanding the background context even more, Turabi and Saynisch (2014) describe the this background coherently and concisely in their paper, and so does the original article by Bhattacharya and Bundorf (2010).

[9] In this paper, B&B will only refer to the paper of by Bhattacharya and Bundorf (2010), not as Bed and Breakfast, perhaps unfortunately for the travel-liking reader.

Finding Impact of Obesity on Wages: Genetic Matching

To improve the internal and inferential validity of the estimated effects of interest, I applied genetic matching instead of Mahalanobis Distance matching and instead of no-matching at all, then checked the estimated effect between the matched groups and its confidence interval.
In order to: (1) inspect the validity and relevance for subdividing the data according to race in gender such as was done in T&S without rigorous justification; and (2) determine which covariates would be possible to not match on, because their unusually large number defeats our ability to match well on all covariates.

Examining Correlation of Obeisty and Wage to Other Variables

I ran linear regression of obesity on all covaraites, and of wage on all covariates, to have a rough idea of the relation and correlation type. Results are detailed in the Appendix.
I observed that from the varaibles highest correlated with obesity and wages which our relevant to this study topic, race and gender are indeed amongst the highest correlated and thus relevant. Moreover, along with obesity, they serve as potentially the most detrimental attributes employers can tell about their employees firstly before deciding on their wage, and thus serve a fruitful ground for discrimination. While differences in other parameters actually can have a valid justification for a higher wage, like higher education or living and working in the city versus rural areas, the employee's appearance should not, but evidently it is associated with different wages and therefore it is wise to inspect the differences individually.
For the reasons above, I roughly agreed with the subdivision by race and gender of T&S, except for one factor: omitting analysis of "Other" races. The responders data is categorized into races "White", "Black", or "Other" – which contains a relatively small amout of responders, but they still should not be excluded from the study as they were in T&S. This limits the scope of generalizability and external validity of their study, to be irrelevant for "Other" races. However, we shuold not discriminate from studying them. From the mentioned regressions, we see that obesity is as related to being of an "Other" race as it is to being "Black"; it had the closest coefficient estimate to "Black".
This suggests that being any race other than white is similarly related with obesity. As for wages, "Other" races had the 3rd closest coefficient estimate (in real value, not absolute) to "Black" out of 34 covariates. Therefore, it seems that the natural division for two groups by race should have been to "White" versus "Nonwhite". Moreover, considering there are under 10% "Other" raced responders as there are Black, and their treatment and outcomes are correlated, it would seem to not skew the estimation too much. Hence I exmained Nonwhite Women first. However, T&S describe the outcomes for Black Men as surpsigingly different than before; therefore, when arriving to check for nonwhite men, I instead adhered to the same initial sample as T&S and checked only Black Men, to compare results more directly. If one is interested in either sample of Nonwhite Men or Black Women, you may see full process and code for further answers.

Coefficient Estimates:	Obesity	Wage
Race: Black	0.39	-1.09
Race: Other	0.30	-0.65

see full table at end.

Secondly, I used the correlation from this regression to consider which variables are possilbe to drop from matching, if anything.

Genetic Matching

Another difference from T&S study was *when* we matched. T&S matched the entire sample first, then divided it into groups. However, the groups were not well balanced on race and gender before division: Post-Macthing Black covariate was 0.15 for non-obese and 0.17 for obese, with the matching p-value under <0.001. This means that after dividing the data, the respective obese vs. non-obese groups are even less balanced then before. Therefore, I *first divided the data into gender and race pairs stratas,* and *then* applied genetic matching individually for obese versus non-obese groups within each strata.

I used the generalized linear model on treatment (obesity) to obtain Propensity Scores – an indicator of the aggregate liklihood of the unit (person) to be in the treatment group. These are sometimes used as a sole metric for matching; however it is one-dimentional and would loose information depth about the distance from each covariate. Therefore, it is best to attach the propensity score for each observation as another covariate, and perform Genetic Matching on them as well as the other covariates; this is what I applied here. I inspected balance before and after Genetic Matching run for 200 generations and population size. As we have a huge dataset with ~4000-9000 responders in each strata, we can afford to find near-exact matches and drop unoptimally matched units. Hence I limited caliper size (the maximum distance of a matched unit from the original) to 0.2, as I found it optimal amongst the calipers examined in improving balance on key covariates while not dropping too many observations.

Appendix II – Exhibits

*For all exhibits, data, process and replication, please see the presentable Web and R Markdown created files in this Shared Link.

Below I present the mentioned exhibits in the text first:

IMPACT OF INSURANCE ON WAGE

Sensitivity Analysis
For Obese:

Sensitivity analysis results:
Rosenbaum Sensitivity Test for Hodges-Lehmann Point Estimate
Unconfounded estimate 3.33

Gamma	Lower bound	Upper bound
1.0	3.33	3.33
1.1	3.03	3.63
1.2	2.73	3.83
1.3	2.53	4.03
1.4	2.33	4.33
1.5	2.13	4.53

For non-obese:

Sensitivity analysis results:
Rosenbaum Sensitivity Test for Hodges-Lehmann Point Estimate
Unconfounded estimate 3.5649

Gamma	Lower bound	Upper bound
1.0	3.5649	3.5649
1.1	3.1649	3.8649
1.2	2.9649	4.1649
1.3	2.7649	4.3649
1.4	2.5649	4.5649
1.5	2.3649	4.7649

QUNATILE EFFECTS

The 5% quantile effects of obesity on wage for white women, calculated by :

```
quants_treat <- quantile(Y1, probs = seq(0,1,0.05))
quants_control <- quantile(Y0, probs = seq(0,1,0.05))
quants_treat-quants_control    are:
       0%          5%         10%         15%         20%         25%
  2.5000000  -0.8059998  -0.8600001  -0.9870001  -1.4760004  -1.3150001
        30%         35%         40%         45%         50%         55%
 -1.2320002  -1.1749999  -1.0119995  -1.0540001  -1.1400003  -1.1579999
```

```
       60%        65%        70%        75%        80%        85%
 -1.2660002 -1.3269999 -0.9319998 -1.6150002 -1.7580004 -2.0620003
       90%        95%       100%
 -1.7800003 -2.2680002 165.0000000
They are all negitive excpet for the data outliers.
```

DICTIONARY OF USED VARIABLES

Variable Coding	Variable Meaning
[1,] "CPS_hourly_rec"	"Hourly wage*"
[2,] "d_hinsEMP"	"Employer coverage in own name"
[3,] "d_hinsNONE"	"Uninsured"
[4,] "d_obese"	"Obese (BMI>30)"
[5,] "d_obese1"	"Mildly obese (BMI>30 and BMI<35)"
[6,] "d_obese2"	"Morbidly obese (BMI>35)"
[7,] "d_overwt"	"Overweight"
[8,] "d_obesinsEMP"	"Obese * employer coverage (own)"
[9,] "d_sexf"	"Female"
[10,] "childany"	"Any children in household"
[11,] "childf"	"Female with children in household"
[12,] "d_race_b"	"Race - Black"
[13,] "d_race_o"	"Race - Other"
[14,] "d_marrnever"	"Never married"
[15,] "d_marroth"	"Formerly married"
[16,] "age"	"Age"
[17,] "d_educ9_12"	"Education: 9-12"
[18,] "d_educ13_up"	"Education: 13 and over"
[19,] "d_AFQT_0_25"	"AFQT: 0-25"
[20,] "d_AFQT_25_50"	"AFQT: 25-50"
[21,] "d_AFQT_50_75"	"AFQT: 50-75"
[22,] "d_AFQT_75_100"	"AFQT: 75-100"
[23,] "d_urban_res"	"Urban residence"
[24,] "d_tenure0_1"	"Job tenure: 0-1 years"
[25,] "d_tenure1_3"	"Job tenure: 1-3 years"
[26,] "d_tenure3_6"	"Job tenure: 3-6 years"
[27,] "d_tenure6_up"	"Job tenure: 6+ years"
[28,] "d_emp0_9"	"Employer size: 0-9"
[29,] "d_emp10_24"	"Employer size: 10-24"
[30,] "d_emp25_49"	"Employer size: 25-49"
[31,] "d_emp50_999"	"Employer size: 50-999"

```
[32,] "d_emp1000_up"    "Employer size: 1000+"
[33,] "d_year1989"      "Survey year: 1989"
[34,] "d_year1990"      "Survey year: 1990"
[35,] "d_year1992"      "Survey year: 1992"
[36,] "d_year1993"      "Survey year: 1993"
[37,] "d_year1994"      "Survey year: 1994"
[38,] "d_year1996"      "Survey year: 1996"
[39,] "d_year1998"      "Survey year: 1998"
[40,] "d_year2000"      "Survey year: 2000"
[41,] "d_year2002"     "Survey year: 2002"
[42,] "d_ind_ag"      "Industry: Agriculture"
[43,] "d_ind_for"      "Industry: Forestry"
[44,] "d_ind_mining"    "Industry: Mining"
[45,] "d_ind_const"     "Industry: Construction"
[46,] "d_ind_mfrg"      "Industry: Manufacturing"
[47,] "d_ind_transp"    "Industry: Transport"
[48,] "d_ind_wtrade"    "Industry: Wholesale trade"
[49,] "d_ind_rtrade"    "Industry: Retail trade"
[50,] "d_ind_finance"   "Industry: Finance"
[51,] "d_ind_bus_svc"   "Industry: Business services"
[52,] "d_ind_pers_svc"  "Industry: Personal services"
[53,] "d_ind_entert"    "Industry: Entertainment"
[54,] "d_ind_prof_svc"  "Industry: Professional services"
[55,] "d_ind_pub_ad"    "Industry: Public administration"
[56,] "d_occ_mgmt"      "Occupation: Management"
[57,] "d_occ_tech"      "Occupation: Technical"
[58,] "d_occ_admin"     "Occupation: Administrative"
[59,] "d_occ_svc"       "Occupation: Service"
[60,] "d_occ_farming"   "Occupation: Farming"
[61,] "d_occ_prodxn"    "Occupation: Production"
[62,] "d_occ_operators" "Occupation: Operators"
[63,] "d_occ_military"  "Occupation: Military"
[64,] "BMI"             "BMI"
```

CORRELATION OF FACTORS WITH OBESITY

Generalized Linear Model Results

I fitted a general linear model of obesity onto all the covaraites, to see which covariates are most highly correlated.

Coefficients:

```
            Estimate Std. Error z value Pr(>|z|)
```

(Intercept)	-1.170e+02	1.486e+01	-7.870	3.54e-15	***
d_sexf	2.581e-01	4.949e-02	5.215	1.84e-07	***
childany	2.482e-01	4.408e-02	5.632	1.78e-08	***
childf	-3.515e-01	6.260e-02	-5.615	1.96e-08	***
age	3.819e-02	6.788e-03	5.626	1.84e-08	***
d_urban_res	-1.957e-01	3.430e-02	-5.705	1.16e-08	***
srvy_yr	5.766e-02	7.546e-03	7.641	2.15e-14	***
AFQTrevised	-3.449e-04	7.537e-04	-0.458	0.647241	
educ	-7.148e-02	8.860e-03	-8.068	7.14e-16	***
tenure	2.096e-04	5.881e-05	3.565	0.000364	***
NumberEmp	1.102e-09	1.268e-06	0.001	0.999306	
d_race_b	3.977e-01	3.672e-02	10.830	< 2e-16	***
d_race_o	3.047e-01	6.562e-02	4.644	3.42e-06	***
d_marrnever	1.474e-01	4.129e-02	3.569	0.000358	***
d_marroth	-2.578e-01	4.252e-02	-6.063	1.34e-09	***
d_ind_ag	3.634e-02	1.939e-01	0.187	0.851359	
d_ind_for	1.281e+00	5.035e-01	2.544	0.010968	*
d_ind_mining	-1.636e-01	2.287e-01	-0.716	0.474282	
d_ind_const	-3.599e-01	1.627e-01	-2.213	0.026926	*
d_ind_mfrg	-1.038e-02	1.534e-01	-0.068	0.946048	
d_ind_transp	2.348e-02	1.578e-01	0.149	0.881703	
d_ind_wtrade	6.121e-02	1.673e-01	0.366	0.714494	
d_ind_rtrade	1.525e-01	1.555e-01	0.981	0.326580	
d_ind_finance	-1.897e-01	1.614e-01	-1.176	0.239726	
d_ind_bus_svc	6.314e-02	1.590e-01	0.397	0.691265	
d_ind_pers_svc	1.533e-01	1.742e-01	0.880	0.378804	
d_ind_entert	-2.231e-01	2.125e-01	-1.050	0.293639	
d_ind_prof_svc	1.542e-01	1.549e-01	0.996	0.319278	
d_occ_mgmt	-1.346e-01	1.410e-01	-0.954	0.340029	
d_occ_tech	-6.343e-02	1.437e-01	-0.441	0.658925	
d_occ_admin	6.097e-02	1.421e-01	0.429	0.667910	
d_occ_svc	8.527e-02	1.429e-01	0.597	0.550676	
d_occ_farming	-1.443e-01	1.773e-01	-0.814	0.415788	
d_occ_prodxn	4.246e-02	1.400e-01	0.303	0.761677	
d_occ_operators	3.643e-03	1.381e-01	0.026	0.978950	

CORRELATION OF FACTORS WITH WAGE

We show both raw results from regression, and later the beta weights sorted from smallest (absolute size) to largest.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1.073e+03	8.117e+01	-13.220	< 2e-16	***
d_sexf	**-1.321e+00**	2.623e-01	-5.037	4.75e-07	***
childany	1.728e+00	2.367e-01	7.300	2.95e-13	***

```
childf         -1.945e+00  3.392e-01  -5.736 9.79e-09 ***
age             4.015e-02  3.703e-02   1.084 0.278275
d_urban_res     1.363e+00  1.928e-01   7.067 1.62e-12 ***
srvy_yr         5.349e-01  4.122e-02  12.976  < 2e-16 ***
AFQTrevised     5.341e-02  4.039e-03  13.224  < 2e-16 ***
educ            8.434e-01  4.789e-02  17.611  < 2e-16 ***
tenure          3.916e-03  3.429e-04  11.421  < 2e-16 ***
NumberEmp       6.698e-06  6.736e-06   0.994 0.320043
d_race_b       -1.099e+00  2.055e-01  -5.346 9.06e-08 ***
d_race_o       -1.560e-01  3.719e-01  -0.419 0.674883
d_marrnever    -9.653e-01  2.242e-01  -4.305 1.67e-05 ***
d_marroth      -5.377e-01  2.274e-01  -2.364 0.018073 *
d_ind_ag       -2.585e+00  7.191e-01  -3.595 0.000325 ***
d_ind_for      -7.583e-01  3.195e+00  -0.237 0.812377
d_ind_mining    5.478e-01  9.246e-01   0.592 0.553553
d_ind_const     2.276e+00  3.943e-01   5.772 7.93e-09 ***
d_ind_mfrg      1.077e+00  2.781e-01   3.873 0.000108 ***
d_ind_transp    2.540e+00  3.458e-01   7.347 2.08e-13 ***
d_ind_wtrade    8.286e-02  4.582e-01   0.181 0.856487
d_ind_rtrade   -1.736e+00  2.974e-01  -5.836 5.39e-09 ***
d_ind_finance   2.307e+00  3.502e-01   6.587 4.57e-11 ***
d_ind_bus_svc   8.810e-01  3.502e-01   2.515 0.011893 *
d_ind_pers_svc -1.527e+00  5.336e-01  -2.862 0.004206 **
d_ind_entert   -1.285e+00  7.746e-01  -1.659 0.097107 .
d_occ_mgmt      6.163e+00  8.933e-01   6.899 5.34e-12 ***
d_occ_tech      4.210e+00  9.049e-01   4.652 3.30e-06 ***
d_occ_admin     2.235e+00  9.005e-01   2.482 0.013081 *
d_occ_svc       1.770e+00  9.069e-01   1.952 0.050955 .
d_occ_farming   3.602e+00  1.083e+00   3.326 0.000883 ***
d_occ_prodxn    3.169e+00  8.916e-01   3.555 0.000379 ***
d_occ_operators 2.175e+00  8.824e-01   2.465 0.013711 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.81 on 31142 degrees of freedom
Multiple R-squared:  0.155,  Adjusted R-squared:  0.1541
F-statistic: 173.1 on 33 and 31142 DF,  p-value: < 2.2e-16
```

Correlation with Wage – Sorted List of Betas

Ordered from the least correlated (to Drop later) to the most:

```
NumberEmp :  0.00000669
tenure :  0.003915773
age :  0.04014622
AFQTrevised :  0.05340964
d_ind_wtrade :  0.08285885
srvy_yr :  0.5348925
d_marroth :  0.5376605
d_ind_mining :  0.5477857
d_race_o :  0.6560026
d_ind_for :  0.7582588
educ :  0.8433695
d_ind_bus_svc :  0.881009
d_marrnever :  0.9653066
d_ind_mfrg :  1.077095
d_race_b :  1.098689
d_ind_entert :  1.28519
d_sexf :  1.321021
d_urban_res :  1.362609
d_ind_pers_svc :  1.527283
childany :  1.727576
d_ind_rtrade :  1.735784
d_occ_svc :  1.770294
childf :  1.94544
d_occ_operators :  2.17507
d_occ_admin :  2.23467
d_ind_const :  2.275741
d_ind_finance :  2.306869
d_ind_transp :  2.540484
d_ind_ag :  2.584878
d_occ_prodxn :  3.169206
d_occ_farming :  3.602088
d_occ_tech :  4.209769
d_occ_mgmt :  6.162826
`(Intercept)` :  1073.065
```

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-1.170e+02	1.486e+01	-7.870	3.54e-15	***
d_sexf	2.581e-01	4.949e-02	5.215	1.84e-07	***
childany	2.482e-01	4.408e-02	5.632	1.78e-08	***
childf	-3.515e-01	6.260e-02	-5.615	1.96e-08	***
age	3.819e-02	6.788e-03	5.626	1.84e-08	***

d_urban_res	-1.957e-01	3.430e-02	-5.705	1.16e-08	***
srvy_yr	5.766e-02	7.546e-03	7.641	2.15e-14	***
AFQTrevised	-3.449e-04	7.537e-04	-0.458	0.647241	
educ	-7.148e-02	8.860e-03	-8.068	7.14e-16	***
tenure	2.096e-04	5.881e-05	3.565	0.000364	***
NumberEmp	1.102e-09	1.268e-06	0.001	0.999306	
d_race_b	3.977e-01	3.672e-02	10.830	< 2e-16	***
d_race_o	3.047e-01	6.562e-02	4.644	3.42e-06	***
d_marrnever	1.474e-01	4.129e-02	3.569	0.000358	***
d_marroth	-2.578e-01	4.252e-02	-6.063	1.34e-09	***
d_ind_ag	3.634e-02	1.939e-01	0.187	0.851359	
d_ind_for	1.281e+00	5.035e-01	2.544	0.010968	*
d_ind_mining	-1.636e-01	2.287e-01	-0.716	0.474282	
d_ind_const	-3.599e-01	1.627e-01	-2.213	0.026926	*
d_ind_mfrg	-1.038e-02	1.534e-01	-0.068	0.946048	
d_ind_transp	2.348e-02	1.578e-01	0.149	0.881703	
d_ind_wtrade	6.121e-02	1.673e-01	0.366	0.714494	
d_ind_rtrade	1.525e-01	1.555e-01	0.981	0.326580	
d_ind_finance	-1.897e-01	1.614e-01	-1.176	0.239726	
d_ind_bus_svc	6.314e-02	1.590e-01	0.397	0.691265	
d_ind_pers_svc	1.533e-01	1.742e-01	0.880	0.378804	
d_ind_entert	-2.231e-01	2.125e-01	-1.050	0.293639	
d_ind_prof_svc	1.542e-01	1.549e-01	0.996	0.319278	
d_occ_mgmt	-1.346e-01	1.410e-01	-0.954	0.340029	
d_occ_tech	-6.343e-02	1.437e-01	-0.441	0.658925	
d_occ_admin	6.097e-02	1.421e-01	0.429	0.667910	
d_occ_svc	8.527e-02	1.429e-01	0.597	0.550676	
d_occ_farming	-1.443e-01	1.773e-01	-0.814	0.415788	
d_occ_prodxn	4.246e-02	1.400e-01	0.303	0.761677	
d_occ_operators	3.643e-03	1.381e-01	0.026	0.978950	

P VALUES BEFORE AND AFTER GENETIC MATCHING

variable Name	T-test p-value Before Matching	p-value After Matching
propensities	.... < 2.22e-16	0.13357
d_sexf	 1	1
childany	.... 0.03091	1

childf
.... 0.03091 1

age
.... < 2.22e-16 0.74044

d_urban_res
.... 0.00085953 0.17964

srvy_yr
.... < 2.22e-16 0.77865

educ
.... 6.8268e-05 0.10641

d_race_b
.... 3.7125e-07 0.36573

d_race_o
.... 3.7125e-07 0.36573

d_marrnever
.... 1.8321e-06 0.11954

d_marroth
.... 0.0028595 0.53455

d_ind_ag
.... 0.10766 1

d_ind_for
.... 0.3174 1

d_ind_mining
.... 0.93361 1

d_ind_const
.... 0.58678 1

d_ind_mfrg
.... 0.28725 0.73893

*** To extract this table of only p-values, I made a short python script which might be helpful to others:

```
#open file
f = open('Txt-GenMatchOutput.txt')

#f=raw_input("Insert GenMatch Output")

results = []
names = []
content = f.readlines()
for line in content:
      print line
   if line[0:3] == "***":
      start = line.find("$") +1
      end =  line.rfind (" ")
      print line[start:end]
      names.append(line[start:end])
   elif line.__contains__("T-test p-value"):
      print "               ", line[(line.find("........")):]

f.close()
```