1a)

This research paper by Schofield seeks to address whether increased caloric intake can result in productivity gains for wage laborers in India.

To answer this question, she conducts 2 studies. The first, a randomized control trial, draws on a sample of rickshaw drivers, many of them malnourished. The randomly assigned treatment group is offered a bundle of cash and 700 calories in food. The control is offered the cash equivalent of this bundle. Differences in BMI, cognitive, and real tasks, are measured afterwards. This experiment yields a significant increase in earnings and labor supply at the 5 and 10 % confidence levels. Gains are greatest for those who are under BMI 20, and below median earners.

In the second design, the author cleverly exploits a quasi-experimental design in assessing the productivity decline associated with the month of fasting during Ramadan, and attributes these declines to decreased caloric intake. Using a triple difference approach, and interacted variables, and even overlapping periods of agricultural seasons and Ramadan as a clever way to finding natural interactions, the author comes to some very methodical conclusions. Declines in productivity are certainly greatest for districts with a higher than median population of Muslims, and I am convinced this decline is owing to Ramadan, and even partly owing to the caloric decline. I am less convinced by this study, mainly because it is difficult to attribute the loss of productivity solely to the decreased caloric intake during Ramadan, when other cultural behaviors are at work.

Link to article:

<https://scholar.harvard.edu/files/hschofield/files/schofield_calories_and_productivity_2014.01.27new_0.pdf>

Schofield, Heather. "The economic costs of low caloric intake: Evidence from India." *Unpublished Manuscript* (2014).

1b)



The above table is a representation of results from the following regression:



Outcome variables: Here, the outcome `l’ is an individual’s earnings, or labor supply.

Independent variables:

T represents an individual’s treatment status:

T=1, if assigned to treatment, and receiving a food + cash bundle

T=0, if assigned to control, and receiving only a cash package.

P is a binary variable, coded to 1 if an individual is observed in the period following the intervention, and 0 if in the period preceding intervention.

The interacted variable TxP is an interaction between the treatment variable, and the pre-post binary. beta\_3 is the coefficient of interest, to estimate the effects of increased caloric intake on earnings, or labor supply.

X is a vector of other individual characteristics, and covariates.

Theta, and gamma, represent fixed effects for the rickshaw stand location, and program enrollment month.

1c)

Coefficients:

1. beta\_3, column (3): Average hours of labor supplied increased by 107% in the second half of the study, (T=1, P=1) for the treatment group, relative to the control group, statistically significant at the 1% significance level (holding all else constant).
2. beta\_3, column (1): Average fraction of days worked increased by 8% in the second half of the study, (T=1, P=1) for the treatment group, relative to the control group, statistically significant at the 1% significance level (holding all else constant).

R-squared:

1. The R-squared on the average days worked=0.24, conditional on working at all, shows a weakly significant relationship between caloric intake and labor supplied.
2. The R-squared on the average earnings =0.15, measured in levels, shows a weakly significant relationship between caloric intake and earnings, indicating that 15% of the variation in earnings can be explained by this model.

Standard error:

1. For beta\_3, column (3) = 0.33, which is less than half the coefficient, of 1.07, therefore making it statistically significant.

2a)



2b)

The following is the table from the regression of income on controls for West, South, Midwest.



The relevant code for the above regression is:

gen west=0

replace west=1 if region==1

gen northeast=0

replace northeast=1 if region==2

gen south=0

replace south=1 if region==3

gen midwest=0

replace midwest=1 if region==4

2c)

Coefficient Interpretation:

We can interpret the coefficients as meaning that average household income decreases significantly by moving out of the North East to any other part of the US. The following are the estimated decreases, keeping all else equal:

Difference between average monthly income per household in the North East, and West, b\_1=$6,264.99, all else equal.

Difference between average monthly income per household in the North East, and South, b\_2=$9,718.88, all else equal.

Difference between average monthly income per household in the North East, and Midwest, b\_3=$7,578.13, all else equal.

The difference in average household incomes is clearly greatest between the North East and the South, whereas it is smallest when compared with the West.

The constant indicates that the average income of a household from the North East is, a=$40,494.62, and is statistically significant at the 5% significance level.

2d)

Yes, the coefficients for the difference in incomes between the 3 regions (West, South, Midwest) is significant for all regions when compared against a baseline of the North East, keeping all else equal.

Stated differently, household in the North East has an average monthly income significantly greater than that in all 3 other regions of the country, controlling for all other factors.

Both the t-statistic of the 3 regions, and p-values, pass the significance test at the 5% level.

b\_1: |t-statistic|=|-2.41|>1.96, p-value= 0.021<0.05 stat. significant, at 95% confidence level

b\_2: |t-statistic|=|-4.01|>1.96, p-value= 0.000<0.05 stat. significant, at 95% confidence level

b\_3: |t-statistic|=|-2.86|>1.96, p-value= 0.007<0.05 stat. significant, at 95% confidence level

3a.



Coefficient Interpretation:

The constant, a, indicates that the energy consumption at the baseline, when green-house gas emissions are 0, is 203.25 Btu’s, keeping all else constant.

According to b, the coefficient, when per capita green-house gas emissions increase by a ton, per capita energy consumption goes up by 5.72 Btu’s, keeping all else constant.

The F-test is passed at the 5% significance level (0.00<0.05), and both a and b are statistically significant at the 95% confidence level.

a: |t-statistic|=|9.65|>1.96, p-value= 0.000<0.05 stat. significant, at 95% confidence level

b: |t-statistic|=|7.92|>1.96, p-value= 0.000<0.05 stat. significant, at 95% confidence level

Both our R-squared and adj R-squared measures are significant.

Adjusted R-squared=0.59, indicating that per capita green house gas usage explains 59% of the variation in per capita energy usage.

3b.



Coefficient Interpretation:

The constant, a, indicates that in the absence of green house gases being released, when toxic emissions are absent, and when nobody is driving at all, energy consumption is estimated to be 99.33 Btu’s.

The first coefficient, b\_1, indicates that for every unit increase in green-house gases (in Btu) per capita released, energy consumed per capita increases by 4.32 tons, keeping all else constant.

The second coefficient, b\_2, indicates that for every unit increase in toxic emissions per capita (in lbs) released, energy consumed per capita increases by 2.53 tons, keeping all else constant.

The third coefficient, b\_3, indicates that for every unit increase in miles driven per capita (per year), energy consumed per capita increases by 0.01 tons, keeping all else constant.

Adj R-squared Interpretation:

Adj R-squared=0.75, which is pretty high for any model, and indicates that 75% of the variation in energy consumed is accounted for by this model.

Only the first two of our b coefficients (green house gases, and toxic emissions) are statistically significant at the 95% confidence level, using both the t-statistic and the p-value.

a: |t-statistic|=|1.33|<1.96, p-value= 0.192>0.05 not statistically significant

b\_1: |t-statistic|=|6.12|>1.96, p-value= 0.000<0.05 stat. significant, at 95% confidence level

b\_2: |t-statistic|=|5.04|>1.96, p-value= 0.000<0.05 stat. significant, at 95% confidence level

b\_3: |t-statistic|=|1.13|<1.96, p-value= 0.265>0.05 not statistically significant

The F-test is satisfied at the 5% significance level, and the R-squared and adj R-squared measures are significant as a result.

The effect of green house gasses on energy consumption estimated in the bivariate model is adjusted downwards in the multivariate specification, from a change of 5.72 to 4.32 tons, per btu of green-house gases, keeping all else equal.

3b.



The second b\* coefficient, for toxic emissions, indicates that given an increase of one standard deviation in toxic emissions (in lbs of toxics released per capita), the level of energy consumption increases by 0.41 standard deviations (in Btu’s of energy per capita) holding all else constant.

This effect is significant at the 95% confidence level, just as the unstandardized version of this variable was.

3d.

In standardized form, of the variables, green-house gases have the greatest effect per standard deviation.

The beta coefficient for green-house gases, b\*=0.58