**1.) Conceptual Questions**

**a.)** Bias of the coefficient would result because of the negative correlation between training sessions attended and reprimands received. The more training sessions that an employee attends, the less reprimands that employee is likely to have. The particular employee would be more aware of the company guidelines and how to act in certain situations, whereas an employee that never attends sessions wouldn’t be that cognizant, leading them to act in a way that isn’t acceptable via company guidelines and that would result in a reprimand. Thus, the *training* coefficient and the *reprimands* coefficient would have a negative (inverse) sign.

**b.)** There could be random measurement error because parents don’t know how long exactly their child is watching TV. They may think that their child has watched 2 hours of TV, when in fact, the child was home alone for an hour and has watched 3-4 hours of TV. The exact amount of TV watched is unknown unless a timer is put to it. This would then result in an underestimate of the measurement because in most instances, children will under-report to their parents how long they watched TV. The sign of the bias would therefore be negative because *tvhours* and *grade* are probably correlated and *tv hours* probably decrease *grade*.

**c.)** For the father’s answer to be used as an instrument for the mother’s, the father’s response must be able to explain the mother’s answer and the father’s answer must be uncorrelated with the error term. I think this could be used as a valid instrument assuming that both parents are involved in the child’s life (i.e. both parents are living with the child) and assuming the answer’s are independent of one another (the father doesn’t know what the mother said).

**2.) Elections**

**a.)** *voteA* = 43.1712 + 0.0236*expendA*

*N =* 173; *df =* 171; *R^2 =* .1561; *p-value =* .00000

With every additional increase in campaign expenditure by candidate A (in $1000’s), candidate A wins .0236% more of the vote.

**b.)** *voteA* = 49.619 + 0.038*expendA* - .036*expendB*

*N =* 173; *df =* 171; *R^2 =* .5299; *p-value =* .00000

𝜷1HAT has increased, which shows that including the variable that was previously omitted eliminates the negative bias that it had. In the previous equation we didn’t control for the opponent’s expenditure, and including it helps un-bias our estimate.

**c.)** *voteA* = 16.965 – 0.032*expendA* - 0.030*expendB +* 9.382*democA +* .0555*prtystrA*

*N =* 173; *df =* 171; *R^2 =* .6264; *p-value =* .00000

**d.)** Another variable that would be added for control would be the historical data of that particular candidate’s seat (i.e. is the candidate an incumbent, how long have they been in office as a US Representative). We can call this variable *term.*

**3.) Subscriptions**

**a.)** *Subscribers* = 242.179 + 3.826*Price* + *1.202*Income

*N =* 854; *df* = 851; *R^2* = .5424; *P-value* = .0000

**b.)**

**c.)** Regression 1: *Price* = 23.328 + 0.898*Labor* + 0.111*Income*

Regression 2: *Subscribers =* 457.842 – 2.909*­­pricehat* – 1.992*Income*

**d.)** *Subscribers =* 457.842 – 2.909*­­pricehat* – 1.992*Income*