14.662, Spring 2015: Problem Set 3
Due Wednesday 22 April (before class)
Heidi L. Williams
TA: Peter Hull

1 Roy model: Chiswick (1978) and Borjas (1987)

Chiswick (1978) is interested in estimating regressions like the following for a single Census year T (in his case, $T = 1970$):

$$\ln(wage_i(T)) \quad = \quad \mathbf{X_i'}\theta + \delta I_i + \alpha_1 I_i Years_i + \alpha_2 I_i Years_i^2 + \beta_1 I_i Arrive_i + \beta_2 I_i Arrive_i^2 + \epsilon_i,$$

where I_i is an indicator for foreign-born, $Years_i$ counts the number of years since migration, and $Arrive_i$ is the calendar year of arrival.

1. *(6 points) By substituting $Arrive_i = T - Years_i$ into the above regression equation, show mathematically that in a single cross-section β_1, β_2, δ, α_1, and α_2 cannot be separately identified.*
   We have

$$\begin{aligned}\ln(wage_i(T)) &= \mathbf{X_i'}\theta + \delta I_i + \alpha_1 I_i \text{Years}_i + \alpha_2 I_i \text{Years}_i^2 + \beta_1 I_i (T - \text{Years}_i) + \beta_2 I_i (T - \text{Years}_i)^2 + \epsilon_i \\ &= \mathbf{X_i'}\theta + (\delta + \beta_1 T + \beta_2 T^2) I_i + (\alpha_1 - \beta_1 - 2\beta_2 T) I_i \text{Years}_i + (\alpha_2 + \beta_2) I_i \text{Years}_i^2 + \epsilon_i\end{aligned}$$

   In a single cross-section T is constant; a regression of $\ln(wage_i(T))$ on I_i, $I_i\text{Years}_i$, and $I_i\text{Years}_i^2$, controlling for $\mathbf{X_i'}$, then identifies

$$\begin{aligned}\gamma_1 &= \delta + \beta_1 T + \beta_2 T^2 \\ \gamma_2 &= \alpha_1 - \beta_1 - 2\beta_2 T \\ \gamma_3 &= \alpha_2 + \beta_2\end{aligned}$$

   but not any of the individual δ, α_1, α_2, β_1, or β_2 coefficients separately.

2. *(8 points) Re-write your new regression equation from part (1) to let γ_1 represent the coefficient on I_i, γ_2 represent the coefficient on $I_i Years_i$, and γ_3 represent the coefficient on $I_i Years_i^2$. What is $\frac{\partial \gamma_1}{\partial T}$, $\frac{\partial \gamma_2}{\partial T}$, and $\frac{\partial \gamma_3}{\partial T}$? Use these expressions to show that with two years of Census data (say, $T = 1970, 1980$) it is possible to identify β_1, β_2, δ, α_1, and α_2.*
   We have

$$\begin{aligned}\frac{\partial \gamma_1}{\partial T} &= \beta_1 + 2\beta_2 T \\ \frac{\partial \gamma_2}{\partial T} &= -2\beta_2 \\ \frac{\partial \gamma_3}{\partial T} &= 0\end{aligned}$$

   With two years of Census data $\{T_1, T_2\}$, we can run the panel (equivalently, first-differenced) specification

$$\Delta \ln(wage_i) = \rho_0 + \rho_1 I_i + \rho_2 I_i \text{Years}_i + \rho_3 I_i \text{Years}_i^2 + \nu_i$$

   By the above result we know such a specification will allow us to recover the underlying "structural" coefficients. We expect:

$$
\begin{aligned}
&\widehat{\rho_0} \xrightarrow{p} 0 \\
&\widehat{\rho_1} \xrightarrow{p} (\beta_1 + 2\beta_2 T_1)\Delta T \\
&\widehat{\rho_2} \xrightarrow{p} -2\beta_2 \Delta T \\
&\widehat{\rho_3} \xrightarrow{p} 0
\end{aligned}
$$

So that from estimation of the first-differenced and levels specification we can consistently estimate

$$
\begin{aligned}
\beta_1 &= p\lim((\widehat{\rho_1} + \widehat{\rho_2} T_1)/\Delta T) \\
\beta_2 &= p\lim(-\widehat{\rho_2}/(2\Delta T))
\end{aligned}
$$

and

$$
\begin{aligned}
\alpha_1 &= p\lim(\widehat{\gamma_2} + \widehat{\rho}_1 \Delta T) \\
\alpha_2 &= p\lim(\widehat{\gamma_3} + \widehat{\rho}_2/(2\Delta T)) \\
\delta &= p\lim(\widehat{\gamma_1} - (\widehat{\rho_1} + \widehat{\rho}_2 T_1)T_1/\Delta T + \widehat{\rho}_2 T_1^2/(2\Delta T))
\end{aligned}
$$

Inference on these coefficients is easy if both regressions are run simultaneously on duplicated data (this is a useful trick whenever you need to compute statistics off of regression coefficients). Note that this procedure also produces two testable implications of the model; that is, we can test whether we indeed have $\rho_0 = \rho_3 = 0$.

3. *(6 points) In order to identify both the assimilation effect and the cohort indicators while also controlling for Census year indicators, Borjas (1987) imposed the restriction that time-specific shocks have the same effect on log earnings of natives and immigrants. How might you assess the validity of this restriction?* With additional years of data we are overidentified for α_1, α_2, β_1, β_2, and δ, and we can test whether pairwise comparisons of years generate significantly different estimates to test the model's restrictions. Formally, we can rewrite the above equations in terms of regression moments, form the optimal second-step GMM weighting matrix, and compute the usual Hansen J test statistic; in this case this statistic will be distributed approximately chi-squared with $5(N-2)$ degrees of freedom for $N > 2$ Census years. Note that the year of arrival was not added to the Census until the 1970 release, so it would not have been possible for Chiswick to form this test (though now we can with more recent Census data).

2 Roy model: Rothschild and Scheuer (2013)

Rothschild and Scheuer (2013) characterize optimal taxation in a Roy model where individuals can self-select into one of multiple sectors based on relative potential skill. In this problem, you will use data from the Current Population Survey to replicate some of their results.

Consider an economy with a unit mass of individuals who can choose between working in either of two sectors. Each person has a two-dimensional skill endowment $(\theta, \varphi) \in \Theta \times \Phi$. The parameter θ captures an individual's productivity in the Θ-sector and φ captures her ability in the Φ-sector. These endowments are jointly distributed with CDF $F(\theta, \varphi)$. Let $S(\theta, \varphi) \in \{\Theta, \Phi\}$ denote a worker's chosen sector and $P_\Theta = \{(\theta, \varphi) \mid S(\theta, \varphi) = \Theta\}$ denote the set of types who choose the Θ-sector.

Individuals have preferences over consumption c and effort e given by

$$U(c, e) = c - \left(\frac{\epsilon}{1+\epsilon}\right) e^{\frac{1+\epsilon}{\epsilon}}$$

Aggregate effort in the Θ-sector is given by

$$E_\Theta \equiv \int_{P_\Theta} \theta e(\theta, \varphi)\, dF(\theta, \varphi)$$

for effort $e(\theta, \varphi)$, and likewise for aggregate effort in the Φ sector, E_Φ. Output is a Cobb-Douglas function of these aggregate effort levels.

$$Y = E_\Theta^\alpha E_\Phi^{1-\alpha}$$

for $\alpha \in (0, 1)$. Let $E \equiv E_\Theta / E_\Phi$ denote relative aggregate effort.

1. *(6 points) What simplifying assumptions are embedded in the functional form for preferences? In particular, what does the parameter ϵ capture? Use a short derivation from your undergrad micro days to justify your interpretation.*

   Preferences are assumed quasilinear, so there will be no income effects. The disutility of effort is assumed constant across workers and sectors and isoelastic, with the parameter ϵ capturing the elasticity of labor supply. To see this, consider an agent choosing optimal labor supply when facing wages w and a linear budget constraint $c = we$:

$$\max_e we - \left(\frac{\epsilon}{1+\epsilon}\right) e^{\frac{1+\epsilon}{\epsilon}}$$
$$\implies e^* = w^\epsilon$$
$$\frac{\partial \ln e^*}{\partial \ln w} = \epsilon$$

2. *(6 points) Assuming that effort is directly observed by employers, derive an expression for the wage of type (θ, φ) as a function of the equilibrium value of E. Use this result to argue that wages are invariant to the scale of θ and φ.*

   The marginal product of aggregate effort in each sector is

$$\frac{\partial Y}{\partial E_\Theta} = \alpha \left(\frac{E_\Phi}{E_\Theta}\right)^{1-\alpha}$$
$$\frac{\partial Y}{\partial E_\Phi} = (1-\alpha) \left(\frac{E_\Phi}{E_\Theta}\right)^{\alpha}$$

With effort directly observed, the prevailing wage will be the marginal product of a unit of effective effort:

$$\tilde{w}(\theta,\varphi)=\begin{cases}\alpha\theta E^{\alpha-1}, & S(\theta,\varphi)=\Theta\\ (1-\alpha)\varphi E^{\alpha}, & S(\theta,\varphi)=\Phi\end{cases}$$

A worker will choose the sector paying the highest wage, so that observed wages should satisfy

$$w(\theta,\varphi)=\max\{\alpha\theta E^{\alpha-1},(1-\alpha)\varphi E^{\alpha}\}$$

Note that we can write

$$\tilde{w}(\theta,\varphi)=\begin{cases}\alpha Y\frac{\theta}{E_\Theta}, & S(\theta,\varphi)=\Theta\\ (1-\alpha)Y\frac{\varphi}{E_\Phi}, & S(\theta,\varphi)=\Phi\end{cases}$$

Scaling the distribution of θ scales both θ and E_Θ by the same factor, and since only the ratio θ/E_Θ is relevant to wages, we can without loss pick a normalization where $E_\Theta=\alpha Y$; similarly we can normalize E_Φ, so that the distribution of potential wages coincides with the distribution of potential talent.

3. *(8 points) By your argument above, we will proceed as if the distribution of wages and skills coincide. Assume now that potential skills/wages are drawn from a bivariate lognormal distribution with means μ_θ and μ_φ, variances σ_θ^2 and σ_φ^2, and correlation coefficient ρ. We want to estimate these parameters from the observed distribution of wages. To do so, we will take advantage of a useful fact about the bivariate normal distribution (derived in Basu and Ghosh (1978)):*

Let X and Y be distributed bivariate normal with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and correlation coefficient ρ. Let $Z=\max\{X,Y\}$. Then the density of Z is

$$g(z) = \left(\frac{1}{\sigma_x}\right)\phi\left(\frac{z-\mu_x}{\sigma_x}\right)\Phi\left(\frac{z-\tilde{\mu}_y}{\tilde{\sigma}_y}\right)+\left(\frac{1}{\sigma_y}\right)\phi\left(\frac{z-\mu_y}{\sigma_y}\right)\Phi\left(\frac{z-\tilde{\mu}_x}{\tilde{\sigma}_x}\right)$$

where

$$\begin{aligned}\tilde{\mu}_x &= \begin{cases}\frac{1}{\gamma_x}\left[\mu_x-(1-\gamma_x)\mu_y\right] & \gamma_x\neq 0\\ \mu_x-\mu_y & \gamma_x=0\end{cases}\\ \tilde{\sigma}_x &= \begin{cases}\frac{\sigma_x}{|\gamma_x|}\sqrt{1-\rho^2} & \gamma_x\neq 0\\ \sigma_x\sqrt{1-\rho^2} & \gamma_x=0\end{cases}\end{aligned}$$

and

$$\begin{aligned}\gamma_x &= 1-\rho\left(\frac{\sigma_x}{\sigma_y}\right)\\ \gamma_y &= 1-\rho\left(\frac{\sigma_y}{\sigma_x}\right)\end{aligned}$$

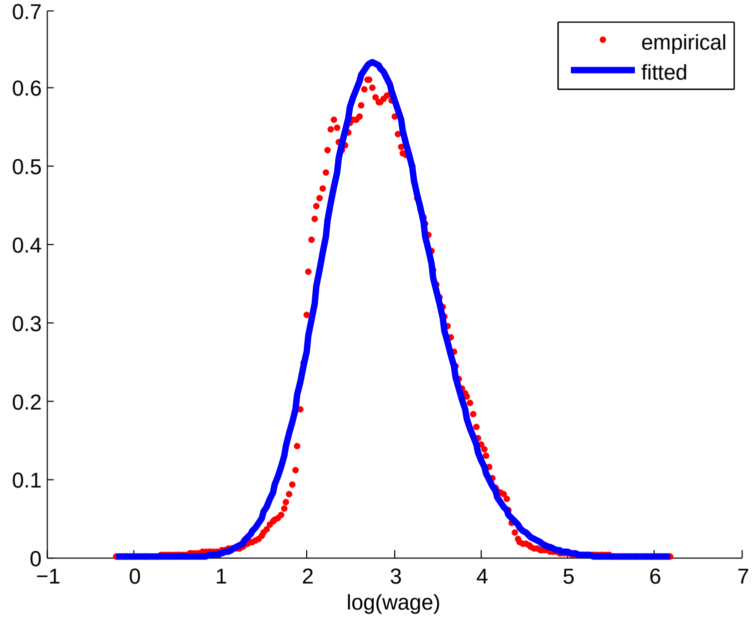
Download the March 2011 CPS earnings and hours data from the NBER website. Generate a sample of log hourly wages from the weekly earnings and weekly hours data.[1] *Please note that there is no extensive margin for labor force participation in this model, so you can restrict your attention to the subset of respondents with positive, non-missing wages. Use the fact above to estimate the parameters of the bivariate wage distribution.*[2] *Use your estimates to generate a predicted wage distribution and plot your prediction against the distribution observed in the CPS.*

My sample has 160,373 observations of log hourly wages who report positive weekly earnings and hours

[1] *The relevant sample weight for the earnings variables is earnwt.*

[2] *You may wish to experiment with different optimization packages and starting values in running this MLE; you might also find that trimming the distribution of raw log wages to drop extreme outliers improves the stability of your estimators.*

Figure 1: Empirical and fitted wage distribution



worked. From this I drop extreme observations of log wages less than zero or above six; this removes less than 0.5% of the sample, leaving a total of 159,651 individuals. Using the above density, the sampling weights, and MATLAB'SÈ *fminsearch* algorithm, I estimate:

$$\mu_\theta = 2.618, \mu_\phi = 2.181$$
$$\sigma_\theta = 0.803, \sigma_\phi = 0.658$$
$$\rho = -0.011$$

These estimates are somewhere in the ballpack of Rothschild and Scheuer (2013), though they appear to be somewhat sensitive to the *fminsearch* starting values. Figure 1 plots the fitted and true distribution of observed wages, which is very similar to the authors' Figure II.

4. *(8 points) Assume that the elasticity of labor supply is 0.5 and that all workers face a marginal tax rate of 0.25 on their wages. Use these values and your estimates from part (3) to determine:*

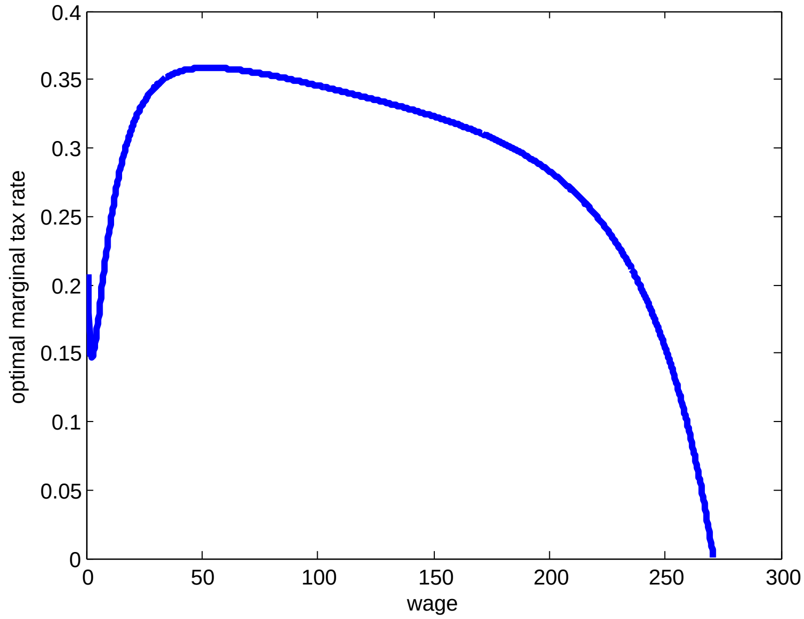
   - *the effort supplied by each worker in her chosen sector*
   - *the share of income paid to each sector*
   - *the parameter α that governs the aggregate production function.*

   Report and interpret your estimate of α here. You do not need to report anything for the first two results; they're simply intermediate steps.
   Individuals facing a marginal tax rate of τ solve

$$\max_e (1-\tau)we - \frac{\epsilon}{1+\epsilon} e^{\frac{1+\epsilon}{\epsilon}}$$
$$\implies e^* = ((1-\tau)w)^\epsilon$$

With $\tau = 0.25$ and $\epsilon = 0.5$ we thus have $e_i^* = \sqrt{0.75 w_i}$. An individual's income is then given by $y = w_i e_i^*$

Figure 2: Optimal tax



and sectoral incomes are

$$Y^{\Theta} = \sum_{P_{\Theta}} \sqrt{0.75} w_i^{3/2}$$
$$Y^{\Phi} = \sum_{P_{\Phi}} \sqrt{0.75} w_i^{3/2}$$

with the parameter α satisfying

$$\alpha = \frac{Y^{\Theta}}{Y^{\Theta} + Y^{\Phi}}$$

Using the parameters estimated above, I draw a sample of 200,000 potential wage pairs and allow individuals to Roy-select into their preferred sector. I then estimate $\alpha = 0.794$. This is somewhat lower than the estimate of 0.942 the authors report, though I am able to replicate this result by simulating potential wages with their estimated parameters.

5. *(8 points) Plot the optimal tax schedule derived by Rothschild and Scheuer (2013) and provided in MTR.mat. Interpret the shape of this schedule. Taking the schedule as given, show how the share of workers in the Θ-sector varies with the sector's offered wage. How does the average effort of Θ-sector workers move with wages? Plot both of these results and interpret.*
The Pareto optimal schedule is plotted in Figure 2. The marginal rate is initially increasing in wages before becoming more regressive beyond around 50 and tending to zero at the extreme of income. More progressive taxation is desirable because it indirectly redistributes income to lower skill workers from higher skill workers, but when there is endogenous selection in the sector higher-skill workers may respond to progressive taxation by shifting to their less productive sector. This effect offsets, so that the optimal marginal tax schedule is more regressive than it would be without Roy selection.

Taking this schedule as given, I calculate optimal effort and earnings for each simulated individual in both sectors and assign the equilibrium sector, observed wage, and effort based on what is individually optimal. The equilibrium share of individuals in the Θ-sector is plotted in Figure 3, while the average effort of Θ-sector workers under the optimal tax schedule is given in Figure 4. Both are monotone but nonlinear in offered wages (and closely resemble the authors' Figure III), with the more workers workers with marginally lower average skill entering the sector for higher wages.

Figure 3: Equilibrium Θ-sector share with optimal tax

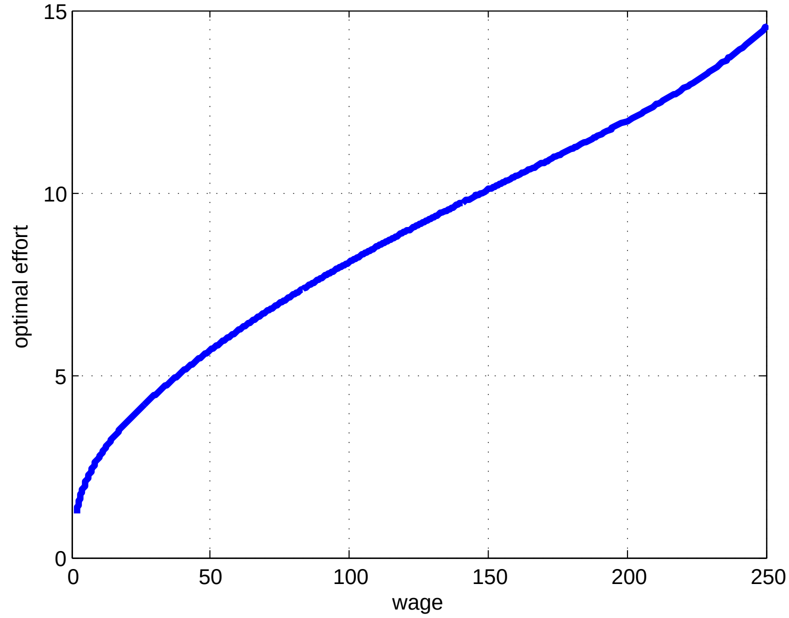
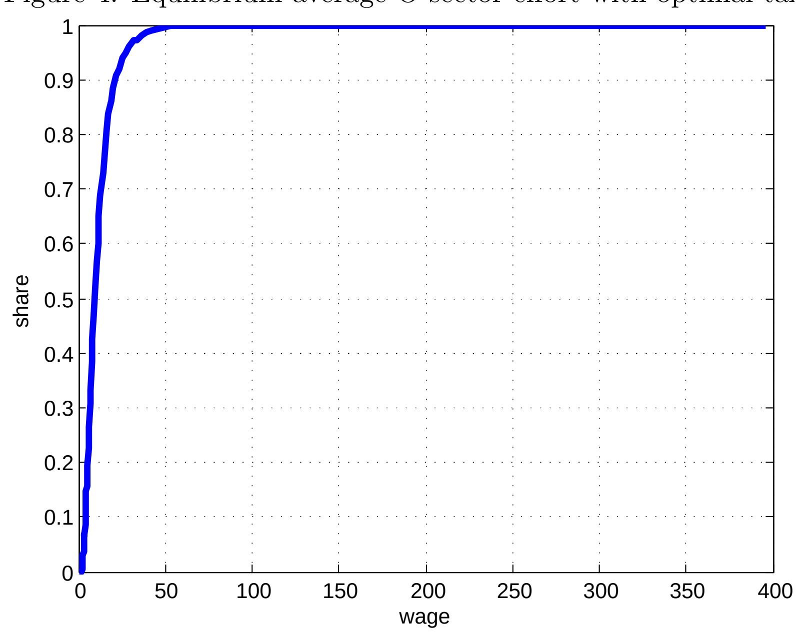


Figure 4: Equilibrium average Θ-sector effort with optimal tax



3 Compensating Differences: Lucas (1977) and Brown (1980)

Suppose true earnings are described by:

$$\ln(wage_{it}) = \beta_0 + \beta_1 Z_{it}^* + \beta_2 X_{it} + \beta_3 A_i + \epsilon_{it}$$

where Z_{it}^* measures working conditions, X_{it} measures observed time-varying worker characteristics, A_i measures unobserved fixed worker characteristics, and ϵ_{it} measures other unobserved factors that affect earnings (such as unmeasured job characteristics). Assume ϵ_{it} is orthogonal to Z_{it}^*, X_{it}, and A_i.

You would like to estimate β_1 - the compensating wage differential paid to workers to offset the disutility of working in jobs with higher levels of the disamenity Z_{it}^*. In practice, you face two estimation problems:

- *"Ability" A_i is unobserved; suppose A_i is negatively correlated with Z_{it}^* conditional on X_{it}*
- *Working conditions Z_{it}^* are measured with error. For example, measurement error could arise if you assigned job characteristics to a survey of workers using a match to the Dictionary of Occupational Titles data based only on occupation and industry, and if that occupation-industry match of job characteristics does not perfectly correspond to characteristics in the worker's specific job. In particular, suppose you observe a noisy measure of working conditions $Z_{it} = Z_{it}^* + \eta_{it}$.*

We'd like to consider the net effect of these two potential sources of bias as well as possible solutions to estimating the compensating differential.

1. *(6 points) Say that you estimate a cross-sectional model as in Lucas (1977):*

$$\ln(wage_{it}) = b_0 + b_1 Z_{it} + b_2 X_{it} + e_{it}$$

For a given t. Suppose in the cross section η_{it} is distributed as independent white noise. Derive an expression for the population regression coefficient b_1 in terms of structural parameters. Can you sign the overall direction of bias?

Denote residuals from the projection of Z_{it} on X_{it} and a constant by $\tilde{Z}_{it}$. We then have

$$\begin{aligned} b_1 &= \frac{Cov(\ln(wage_{it}), \tilde{Z}_{it})}{Var(\tilde{Z}_{it})} \\ &= \frac{Cov(\beta_1 Z_{it}^* + \beta_2 X_{it} + \beta_3 A_i + \epsilon_{it}, \tilde{Z}_{it})}{Var(\tilde{Z}_{it})} \\ &= \frac{\beta_1 Cov(Z_{it}^*, \tilde{Z}_{it}) + \beta_3 Cov(A_i, \tilde{Z}_{it})}{Var(\tilde{Z}_{it})} \end{aligned}$$

Write:

$$
\begin{aligned}
Cov(Z^*_{it}, \tilde{Z}_{it}) =& Var(Z^*_{it}) - \frac{Cov(Z^*_{it}, X_{it})^2}{Var(X_{it})} \\
\equiv& \sigma^2_{Z^*} - \frac{\sigma^2_{Z,^*X}}{\sigma^2_X} \\
Cov(A_i, \tilde{Z}_{it}) =& Cov(A_i, Z^*_{it}) - \frac{Cov(Z^*_{it}, X_{it})Cov(A_i, Z^*_{it})}{Var(X_{it})} \\
\equiv& \sigma_{A,Z^*} - \frac{\sigma_{Z,^*X}\sigma_{A,Z^*}}{\sigma^2_X} \\
Var(\tilde{Z}_{it}) =& Var(Z^*_{it} + \epsilon_{it}) + \frac{Cov(Z^*_{it}, X_{it})^2}{Var(X_{it})} \\
& - 2\frac{Cov(Z^*_{it}, X_{it})}{Var(X_{it})} Cov(Z^*_{it} + \epsilon_{it}, X_{it}) \\
=& Var(Z^*_{it}) + Var(\epsilon_{it}) - \frac{Cov(Z^*_{it}, X_{it})^2}{Var(X_{it})} \\
\equiv& \sigma^2_{Z^*} + \sigma^2_\epsilon - \frac{\sigma^2_{Z,^*X}}{\sigma^2_X}
\end{aligned}
$$

Then

$$
b_1 = \beta_1 \left(\frac{\sigma^2_{Z^*}\sigma^2_X - \sigma^2_{Z,^*X}}{\sigma^2_{Z^*}\sigma^2_X - \sigma^2_{Z,^*X} + \sigma^2_\epsilon \sigma^2_X} \right) + \beta_3 \left(\frac{\sigma^2_X \sigma_{A,Z^*} - \sigma_{Z,^*X}\sigma_{A,Z^*}}{\sigma^2_{Z^*}\sigma^2_X - \sigma^2_{Z,^*X} + \sigma^2_\epsilon \sigma^2_X} \right)
$$

This expression incorporates both omitted-variables bias and attenuation bias in relating the population regression coefficient to the underlying structural parameters. The first term contains a signal-to-noise ratio that captures the attenuation bias due to measurement error in Z_{it}, while the second captures omitted-variables bias, again attenuated by measurement error. We can't sign the total bias without knowing the relationship between $Z_{it}*$, X_{it}, and A_{it}.

2. *(8 points) Say that, like Brown (1980), you find a panel dataset that allows you to estimate a model with individual fixed effects and you estimate*

$$
\Delta \ln(wage_{it}) \quad = \quad b_1 \Delta Z_{it} + b_2 \Delta X_{it} + \Delta e_{it}
$$

Suppose within-individual measurement error is persistent, so that $\eta_{it} = \rho\eta_{it-1} + \nu_{it}$ where ν_{it} is independent (across both time and individuals) white noise. Derive an expression for the population regression coefficient b_1 in terms of structural parameters.

Again partialling out the effect of covariates, we now have the population regression coefficient identifying

$$
\begin{aligned}
b_1 &= \frac{Cov(\Delta \ln(wage_{it}), \Delta \tilde{Z}_{it})}{Var(\Delta \tilde{Z}_{it})} \\
&= \frac{Cov(\beta_1 \Delta Z^*_{it} + \beta_2 \Delta X_{it} + \Delta\epsilon_{it}, \Delta \tilde{Z}_{it})}{Var(\Delta \tilde{Z}^*_{it} + \Delta\eta_{it})} \\
&= \beta_1 \frac{Cov(\Delta Z^*_{it}, \Delta \tilde{Z}_{it})}{Var(\Delta \tilde{Z}_{it})}
\end{aligned}
$$

Writing:

$$\begin{aligned}
Cov(\Delta Z^*_{it}, \Delta\tilde{Z}_{it}) =& Var(\Delta Z^*_{it}) - \frac{Cov(\Delta Z^*_{it}, \Delta X_{it})^2}{Var(\Delta X_{it})} \\
\equiv& \sigma^2_{\Delta Z^*} - \frac{\sigma^2_{\Delta Z^* \Delta X}}{\sigma^2_{\Delta X}} \\
Var(\Delta\tilde{Z}_{it}) =& Var(\Delta Z^*_{it}) + Var(\Delta\epsilon_{it}) - \frac{Cov(\Delta Z^*_{it}, \Delta X_{it})^2}{Var(\Delta X_{it})} \\
=& Var(\Delta Z^*_{it}) + Var((\rho-1)\epsilon_{it-1} + \nu_{it}) - \frac{Cov(\Delta Z^*_{it}, \Delta X_{it})^2}{Var(\Delta X_{it})} \\
=& \sigma^2_{\Delta Z^*} + (\rho-1)^2\sigma^2_\epsilon + \sigma^2_\nu - \frac{\sigma^2_{\Delta Z^* \Delta X}}{\sigma^2_{\Delta X}}
\end{aligned}$$

we have

$$b_1 = \beta_1 \frac{\sigma^2_{\Delta Z^*}\sigma^2_{\Delta X} - \sigma^2_{\Delta Z^* \Delta X}}{\sigma^2_{\Delta Z^*}\sigma^2_{\Delta X} - \sigma^2_{\Delta Z^* \Delta X} + (\rho-1)^2\sigma^2_\epsilon\sigma^2_{\Delta X} + \sigma^2_\nu\sigma^2_{\Delta X}}$$

Since first-differencing removes the bias from omitting individual ability, we are left just with the signal-to-noise formula for attenuation bias, now formulated in changes rather than levels.

3. *(6 points) Briefly discuss what problem(s) are solved by moving to the panel model, and what problem(s) are introduced. Is it always the case that b_1 from the panel model will be an attenuated estimate of β_1, so that we can always consistently estimate the sign of the compensating differential?*
First-differencing leaves us only with attenuation bias; by the Cauchy-Schwartz inequality $\sigma^2_{\Delta Z^*}\sigma^2_{\Delta X} \geq \sigma^2_{\Delta Z^* \Delta X}$ so the ratio multiplying β_1 indeed lies between zero and one. Although this bias may be more severe than in the cross-sectional case (especially if the error in measurement is transitory but the "signal" is persistant), we can at least unambiguously sign β_1 by the sign of b_1.

4 Compensating differences: Gruber and Krueger (1991) and Gruber (1997)

Consider the formalization of the Summers (1989) model from Gruber and Krueger (1991). Suppose that labor demand (L_d) is given by:

$$L_d = f_d(W + C)$$

and suppose labor supply (L_s) is given by:

$$L_s = f_s(W + \alpha C)$$

where C is the cost of mandated health insurance, αC is the monetary value that employees place on health insurance, and W is the wage rate.

1. *(6 points) Derive an expression for how wages change under a mandate ($\frac{dW}{dC}$) in terms of α, the labor demand elasticity η^d, and the labor supply elasticity η^s. Derive an analogous expression for how employment changes under a mandate. Give an intuition for the cases where $\alpha = 0$ and $\alpha = 1$.*
   In equilibrium, $f_s(W^* + \alpha C) = f_d(W + C)$. Taking logs and differentiating with respect to C at $C = 0$ gives

$$\frac{f_s(W^*)}{f_s(W^*)}\left(\frac{\partial W^*}{\partial C} + \alpha\right) = \frac{f_d(W^*)}{f_d(W^*)}\left(\frac{\partial W^*}{\partial C} + 1\right)$$
$$\eta^s \frac{\partial W^*}{\partial C} + \alpha\eta^s = \eta^d \frac{\partial W^*}{\partial C} + \eta^d$$
$$\frac{\partial W^*}{\partial C} = -\left(\frac{\alpha\eta^s - \eta^d}{\eta^s - \eta^d}\right)$$

   Similarly,

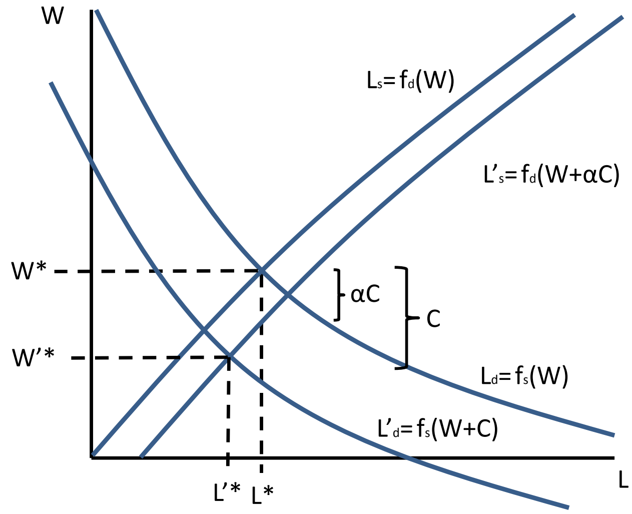
$$\frac{\partial L_d}{\partial C} = f_d(W^*)\left(\frac{\partial W^*}{\partial C} + 1\right)$$
$$= (1-\alpha)\frac{L^*}{W^*}\left(\frac{\eta^s\eta^d}{\eta^s - \eta^d}\right)$$

   Intuitively, if $\alpha = 1$ employees value health insurance as much as it costs to require it, so wages offset the cost one-for-one and total employment is unchanged. As α declines towards zero, the effect on wages diminishes while total employment begins to fall. At $\alpha = 0$ such that employees do not value health insurance at all, wages fall by the least amount and employment declines by the largest amount.

2. *(6 points) Draw a graph of employment (x-axis) against wages (y-axis) with labor supply and labor demand curves before and after the mandated benefit regime. Give some intuition for how to interpret the graph.*
   See Figure 5. A mandated shifts labor demand in by C vertically, reflecting the higher per-unit cost of employing labor. Labor supply also shifts vertically, but by less for $\alpha < 1$, reflecting the employees' willingness to receive a lower wage for some (partial) valuation of the benefit. The new equilibrium following the regime change consists of an unambiguously lower wage that is highest when $\alpha = 0$ and weakly lower total employment that is unchanged when $\alpha = 1$.

3. *(6 points) Describe how the effects of a payroll tax on wages and employment might differ from the effects of a mandated benefit. Would it matter whether the payroll tax collections were used to finance a public health insurance program? What if the public health insurance program had enrollment that was restricted to workers only?*

Figure 5: Labor demand and supply before and after mandated benefits



A payroll tax could differ from a mandated benefit because the revenues could potentially be used to fund programs the valued by employees and non-employees alike. If, say, the revenues were invested in a public health insurance for workers only, the effects could be similar to those of a mandated benefit provided that workers value public and employer-provided health insurance similarly. If both workers and non-workers could take part in public health insurance, the program would have no effect on labor supply because there would be no added benefit to working.

4. *(6 points) Read over the Gruber (1997) paper. He discusses three potential explanations for his results: inelastic labor supply, perfectly elastic labor demand, and full employee valuation of benefits. How might you distinguish between these three potential explanations?*
We could potentially disentangle these results by analyzing various payroll tax and benefit schemes in the same labor market. Suppose we could observe a payroll tax reform used to finance benefits with α plausibly equal to zero (or quite small). The effect of this reform on wages would provide us with a benchmark for $\eta^d/(\eta^s - \eta^d)$, which we could use to rule out the edge case supply and demand explanations.

References

Basu, Asit and Jayanta Ghosh, "Identifiability of the Multinormal and other Distributions under Competing Risks Model," *Journal of Multivariate Analysis*, 1978, *8* (3), 413–429. 3

Borjas, George, "Self-selection and the earnings of immigrants," *American Economic Review*, 1987, *77* (4), 531–553. 3

Brown, Charles, "Equalizing differences in the labor market," *Quarterly Journal of Economics*, 1980, *94* (1), 113–134. 2

Chiswick, Barry, "The effect of Americanization on the earnings of foreign-born men," *Journal of Political Economy*, 1978, *86* (5), 897–921. 1

Gruber, Jonathan, "The incidence of payroll taxation: Evidence from Chile," *Journal of Labor Economics*, 1997, *15* (S3), S72–S101. 4

_ **and Alan Krueger**, "The incidence of mandated employer-provided insurance: Lessons from workers' compensation insurance," in David Bradford, ed., *Tax Policy and the Economy Volume 5*, 1991, pp. 111–143. 4

Lucas, Robert, "Hedonic wage equations and psychic wages in the returns to schooling," *American Economic Review*, 1977, *67* (4), 549–558. 1

Rothschild, Casey and Florian Scheuer, "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 2013, *128* (2), 623–668. 2

Summers, Lawrence, "Some simple economics of mandated benefits," *American Economic Review Papers and Proceedings*, 1989, *79* (2), 177–183. 4