Technical Appendix to
"To Work or Not to Work: Did Tax Reforms Affect Labor Force Participation of Married Couples?"

By Michael Bar* and Oksana Leukhina†

Revised: April, 2007.

Contents

*Bar: Department of Economics, San Francisco State University, San Francisco, CA 94132;
†Leukhina: Department of Economics, University of North Carolina, Chapel Hill, NC 27599.

1 Working with the U.S. Census Data

We download the 1960, 1970, 1980, 1990, 2000 U.S. Census data from IPUMs. Most of the census questions relevant to this project refer to the previous years, i.e. 1959, 1969, ..., 1999. We keep only married non-farm individuals of ages $[25 - 64]$ whose spouse is present and convert all incomes into 1999 dollars using 12 months averages of seasonally adjusted CPI,

1959	1969	1979	1989	1999
29.17	36.68	72.58	123.94	166.58

We then create series that are natural logs of all income types.

We do not correct for topcoding in 89 and 99 because the topcoded observations are already replaced by state mean or median. Hence, we only correct for 59, 69, 79. Using the mean and standard deviation of the truncated distribution of logs of male annual wage inomes, the level of the topcode, and the assumption of the normality of this distribution, we compute the expected mean in the tail of the male wage distribution. The results are reported below.

year	μ_X truncated	σ_X, truncated	topcode: a	correction: $E[X\|X > a]$
1959	10.19074	0.6695618	11.86896571	12.09612871
1969	10.48966	0.6622401	12.33302225	12.53615899
1979	10.46222	0.7789073	12.05602852	12.39249818

We then replace the topcoded male annual wage income with $E[X|X > a]$. We then replace the topcoded female annual wages with a*mean(wage of female)/mean(wage of male) of those individuals whose wage exceeds the mean of male wages and excluding those with topcoded wage income. We deal with topcoded observations of other incomes in the same manner we deal with female wage income.

Once we correct for topcoding we create a new labor income variable,

$$\text{Labor Income} = \text{Wage Income} + \text{Business Income} + \text{Farm Income},$$

and drop individuals with negative labor incomes.

We finally need to deal with intervalled variables. Actual weeks worked last year and usual weekly hours worked last year are available since 1979 only. For 1959 and 1969 we are forced to use intervalled counterparts of these series. The objective is to figure out the right midpoints for each of the intervals. To do so we use 1979 data on actual and intervalled series and compute averages for each interval.

	1959	1969	1979	1989	1999
Actual Hours	NA	NA	Available	Available	Available
Intervalled Hours	Available	Available	Available	Available	NA
Actual Weeks	NA	NA	Available	Available	Available
Intervalled Weeks	Available	Available	Available	Available	Available

We get different midpoints for men and women.

We drop people with a mismatch between hours and income, i.e. positive hours but negative incomes or vice versa. We then match husbands and wives. Here we keep the following variables: year, household weight, personal weight, husband's and wife's labor incomes last year, their hours, age, race, education record, number of children ever born and number of children under five at home, and class of work (whether they are self-employed, work for wage, or neither), and weeks worked last year.

The number of observations (couples) that we end up with is given by

Sample 1: year	# couples
1959	$21,897,992$
1969	$24,218,210$
1979	$34,481,282$
1989	$37,712,472$
1999	$42,328,021$

We drop the no earner couples (both husband and wife work 0 hours and earn 0 income). After this, the number of available observation changes as follows:

Sample 2: year	# couples	fraction of Sample 1 couples (by year)
1959	$21,449,563$	0.020478088
1969	$23,623,713$	0.02454752
1979	$33,093,874$	0.040236555
1989	$36,280,387$	0.037973777
1999	$40,794,924$	0.036219435

2 Solving the Model

The household's problem is $\max\left\{V_{2E}\left(w_m,w_f\right),V_{1M}\left(w_m,w_f\right),V_{1F}\left(w_m,w_f\right)\right\}$, where

$$V_{2E}\left(w_m,w_f\right)=\max_{c_m,c_f}\left\{\lambda\left[\alpha\log\left(c_m\right)+\left(1-\alpha\right)\log\left(1-h_m\right)\right]+\left(1-\lambda\right)\left[\alpha\log\left(c_f\right)+\left(1-\alpha\right)\log\left(1-h_f\right)\right]\right\}$$
$$\text{s.t.}\quad \left(1+\tau\right)\left(c_m+c_f\right)=\left(w_m+w_f\right)\eta-T\left(w_m,w_f\right)$$

and the value functions of 1M and 1F couples are the obvious special cases of the above. That is, V_{1M} is the same as the above with $w_f=h_f=0$, and V_{1F} is the same as above with $w_m=h_m=0$. We use the following notation to denote the after tax effective income

$$\begin{aligned} I_{2E} &\equiv \frac{\left(w_m+w_f\right)\eta-T\left(w_m,w_f\right)}{1+\tau}, \\ I_{1M} &\equiv \frac{w_m\eta-T\left(w_m,0\right)}{1+\tau}, \\ I_{1F} &\equiv \frac{w_f\eta-T\left(0,w_f\right)}{1+\tau}. \end{aligned}$$

where η is the ratio of total annual income to annual wage and salary earnings. We assume that η is the same for all the couples.

Notice that for any given after tax income, a fraction λ will be consumed by the male and a fraction $(1-\lambda)$ will be consumed by the female. Thus, we can write the above maximization problem as

$$\begin{aligned} V_{2E} &= \lambda\left[\alpha\log\left(\lambda I_{2E}\right)+(1-\alpha)\log\left(1-h_m\right)\right] \\ &\quad +(1-\lambda)\left[\alpha\log\left((1-\lambda)I_{2E}\right)+(1-\alpha)\log\left(1-h_f\right)\right] \\ &= \lambda\alpha\log(\lambda)+\lambda\alpha\log\left(I_{2E}\right)+\lambda(1-\alpha)\log\left(1-h_m\right) \\ &\quad +(1-\lambda)\alpha\log(1-\lambda)+(1-\lambda)\alpha\log\left(I_{2E}\right)+(1-\lambda)(1-\alpha)\log\left(1-h_f\right) \\ &= \alpha\log\left(I_{2E}\right)+(1-\alpha)\left[\lambda\log\left(1-h_m\right)+(1-\lambda)\log\left(1-h_f\right)\right]+\psi \end{aligned}$$

Where $\psi=\alpha\left[\lambda\log(\lambda)+(1-\lambda)\log(1-\lambda)\right]$. Thus, the maximum value functions can be reduced to

$$\begin{aligned} V_{2E} &= \alpha\log\left(I_{2E}\right)+(1-\alpha)\left[\lambda\log\left(1-h_m\right)+(1-\lambda)\log\left(1-h_f\right)\right]+\psi \\ V_{1M} &= \alpha\log\left(I_{1M}\right)+(1-\alpha)\lambda\log\left(1-h_m\right)+\psi \\ V_{2E} &= \alpha\log\left(I_{1F}\right)+(1-\lambda)(1-\alpha)\log\left(1-h_f\right)+\psi \end{aligned}$$

In what follows, the constants can be dropped since they do not affect the participation choice. We thus abuse notation and denote the simplified maximum value functions, without the constant term, by the same notation as the original ones

$$\begin{aligned} V_{2E} &= \alpha\log\left(I_{2E}\right)+(1-\alpha)\left[\lambda\log\left(1-h_m\right)+(1-\lambda)\log\left(1-h_f\right)\right] \\ V_{1M} &= \alpha\log\left(I_{1M}\right)+(1-\alpha)\lambda\log\left(1-h_m\right) \\ V_{2E} &= \alpha\log\left(I_{1F}\right)+(1-\alpha)(1-\lambda)\log\left(1-h_f\right) \end{aligned}$$

Decision rules

Our model implies the partition of the wage space into 3 regions: 2E, 1M, and 1F. The typical regions look approximately as shown in Figure 5. In the following sections, we present the thresholds for the calibrated version of the model, and those will be non-linear.

The thresholds between the regions are functions $L(\cdot)$ and $H(\cdot)$ that for all w_m solve

$$\begin{aligned} V_{2E}\left(w_m, L\left(w_m\right)\right) &= V_{1M}\left(w_m, 0\right), \\ V_{2E}\left(w_m, H\left(w_m\right)\right) &= V_{1F}\left(0, H\left(w_m\right)\right). \end{aligned}$$

These functions depend on the tax code and preference parameters. Notice that the parameters of the wage distribution determine where the couples are located in the wage space, while the rest of the model parameters as well as the tax laws determine the shape and location of the two decision rule thresholds.

The household chooses to be a 2E couple rather than a 1M couple if $V_{2E} \geq V_{1M}$, that is

$$\begin{aligned}
\alpha \log \left(I_{2E}\right) + (1-\alpha)\left[\lambda \log \left(1-h_m\right) + (1-\lambda) \log \left(1-h_f\right)\right] &\geq \alpha \log \left(I_{1M}\right) + \lambda (1-\alpha) \log \left(1-h_m\right) \\
\alpha \log \left(I_{2E}\right) + (1-\alpha)(1-\lambda) \log \left(1-h_f\right) &\geq \alpha \log \left(I_{1M}\right) \\
\log \left(\frac{I_{2E}}{I_{1M}}\right) &\geq -\frac{(1-\alpha)(1-\lambda)}{\alpha} \log \left(1-h_f\right) \\
\log \left(\frac{\left(w_m + w_f\right)\eta - T\left(w_m, w_f\right)}{w_m \eta - T\left(w_m, 0\right)}\right) &\geq -\frac{(1-\alpha)(1-\lambda)}{\alpha} \log \left(1-h_f\right) \\
\frac{\left(w_m + w_f\right)\eta - T\left(w_m, w_f\right)}{w_m \eta - T\left(w_m, 0\right)} &\geq \phi \\
w_m \eta + w_f \eta - T\left(w_m, w_f\right) &\geq \phi w_m \eta - \phi T\left(w_m, 0\right) \\
w_m \eta - \phi w_m \eta + w_f \eta &\geq T\left(w_m, w_f\right) - \phi T\left(w_m, 0\right) \\
(1-\phi) w_m \eta + w_f \eta &\geq T\left(w_m, w_f\right) - \phi T\left(w_m, 0\right) \\
\frac{(1-\phi) w_m + w_f}{w_f} &\geq \frac{T\left(w_m, w_f\right) - \phi T\left(w_m, 0\right)}{w_f \eta}
\end{aligned}$$

where $\phi = \exp\left(-\frac{(1-\alpha)(1-\lambda)}{\alpha} \log \left(1-h_f\right)\right)$.

Linear taxes

$$\begin{aligned}
\frac{\left(w_m + w_f\right)\eta - T\left(w_m, w_f\right)}{w_m \eta - T\left(w_m, 0\right)} &\geq \phi \\
\frac{\left(w_m + w_f\right)\eta - t\left(w_m + w_f\right)\eta}{w_m \eta - t w_m \eta} &\geq \phi \\
\frac{(1-t)\left(w_m + w_f\right)\eta}{(1-t) w_m \eta} &\geq \phi \\
\frac{w_m + w_f}{w_m} &\geq \phi
\end{aligned}$$

3 Integration

We describe the methods used for computing the moments in our model. Green 2000 [1] points out that "A longstanding challenge in applied econometrics has been to obtain a fast and accurate method of computing cumulative probabilities for the bivariate normal distribution". We admit that for us it has been a challenge indeed. Our task was to find efficient techniques for computing moments of the form $E\left[g(X,Y) \mid (X,Y) \in A\right]$, where the set A is some nontrivial set in $\mathbb{R}^2$. We start by describing the difficulties in computing the moments in our models and the tradoffs we faced when choosing the integration method. Then, we describe in detail how one can improve the speed of computation by using some basic properties of the truncated normal distribution and appropriate numerical integration techniques.

The tax model implies a partition of the wage space into three regions: 1-earner male, 2-earner, and 1-earner female couples. We would like to compute some moments in the model, conditional on being in particular region. The main difficulty in computing conditional moments in the tax model is the fact that we cannot obtain a closed form solution for the limits of integration, i.e, for the thresholds separating the regions. The reason is that the thresholds are a complicated function of the tax code. The only way of obtaining the thresholds is by using TaxSim tax simulator in order to file taxes for many individuals, solving ther utility maximization problem, and finally approximation of the thresholds by interpolation. Another problem with the tax model is that the tax code introduces non-smoothness of the integrands, especially after the eraned income tax credit was introduced. Fortunately, the model has analytical solution, which enables us to solve the utility maximization problem for many household relatively fast. We face similar difficulty in the model with home production. In this model we cannot obtain analytical solutin of the threshods, and must rely on numerical solution.

We considered several quadrature methods: (1) Gauss-Hermite, (2) Gauss-Legendre, (3) Simpson rule. The Gauss-Hermite quadrature is designed for integrating smooth function over $(-\infty, \infty)$ with respect to the Gaussian distribution. It achieves great accuracy with very few points. We use the Gauss-Hermite quadrature in the computation of open intervals in the model with home production. Since the thresholds are expensive to compute, but they are smoth functions, the Gauss-Hermite quadrature is best suited for those moments. The Gauss-Legendre is efficient for integrating smooth finctions over closed intervals. We implement it in the model with home production for computing conditional probabilities on intervals of husband's income. The Simpson rule is less accurate than the Gaussian quadrature and therfore uses many nodes. It is however superior to the Gaussian quadrature methods when the integrand is highly non-smooth, with sharp kinks. The fact that it is relatively fast to evaluate the integrand in the tax model (due to the existense of analytical solution to the households' problem), and the fact that the integrand in the tax model has kinks, makes the Simpson rule the most suitable for this model. See Miranda 2002 [3], chapter 5, for more detailes and codes for computing nodes and weights for various quadrature methods.

It is highly desirable to reduce double integrals to single integrals if possible. The next section describes how we can do that in our model. We use two basic properties of the bivariate normal distribution.

3.1 Notation

- X, Y are random variables.
- x, y are generic realizations of X and Y respectively.
- $f(x, y)$ is the joint distribution of (X, Y).
- $f_X(x)$ and $f_Y(y)$ are the marginal densities, i.e.,

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\, dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)\, dx \end{aligned}$$

- $f(y|x)$ is the conditional density of Y given $X = x$, i.e.,

$$f(y|x) = \frac{f(x,y)}{f_X(x)}$$

We denote the parameters of conditional distribution by $\mu_{Y|X=x}$ and $\sigma_{Y|X=x}$.

- $f(x|y)$ is the conditional density of X given $Y = y$, i.e.,

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}$$

- $F_X(x)$ and $F_Y(y)$ are the unconditional cumulative density functions (c.d.f.) of X and Y respectively. That is,

$$\begin{aligned} F_X(x) &= \Pr(X \leq x) = \int_{-\infty}^{x} f_X(s)\,ds \\ F_Y(y) &= \Pr(Y \leq y) = \int_{-\infty}^{x} f_Y(s)\,ds \end{aligned}$$

- $F(y|x)$ is the conditional c.d.f. of Y given that $X = x$. i.e.,

$$F(y|x) = \Pr(Y \leq y|X = x) = \int_{-\infty}^{y} f(s|x)\,ds$$

- $F(x|y)$ is the conditional c.d.f. of X given that $Y = y$. i.e.,

$$F(x|y) = \Pr(X \leq x|Y = y) = \int_{-\infty}^{x} f(s|y)\,ds$$

- $\phi(z)$ is the probability density function (p.d.f.) of the standard normal, i.e.,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

- $\Phi(z)$ is the c.d.f. of the standard normal.

3.2 Theorems

Theorem 1 *Let* $(X,Y) \sim N\left(\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}\right)$. *Then*

$$\begin{aligned} (Y|X = x) &\sim N\left[\alpha + \beta x, \sigma_Y^2\left(1 - \rho^2\right)\right] \\ \alpha &= \mu_Y - \beta\mu_X \\ \beta &= \frac{\sigma_{XY}}{\sigma_X^2},\ \rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \end{aligned}$$

Theorem 2 ***(Moments of truncated normal distribution[1]).*** *Let $Y \sim N[\mu, \sigma^2]$. Then*

$$
\begin{aligned}
E[Y|a_1 \le Y \le a_2] &= \mu - \sigma \left[\frac{\phi(\alpha_2) - \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right] \\
E\left[Y^2|a_1 \le Y \le a_2\right] &= \sigma^2 + \mu^2 - \sigma^2 \left[\frac{\alpha_2 \phi(\alpha_2) - \alpha_1 \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right] - 2\mu\sigma \left[\frac{\phi(\alpha_2) - \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right]
\end{aligned}
$$

where

$$\alpha_1 = \frac{a_1 - \mu}{\sigma}, \quad \alpha_2 = \frac{a_2 - \mu}{\sigma}.$$

In what follows we will be interested in computing integrals of the type

$$
\begin{aligned}
\int_{a_1}^{a_2} y f(y)\, dy &= E[Y|a_1 \le Y \le a_2] \cdot [\Phi(\alpha_2) - \Phi(\alpha_1)] \\
\int_{a_1}^{a_2} y^2 f(y)\, dy &= E\left[Y^2|a_1 \le Y \le a_2\right] \cdot [\Phi(\alpha_2) - \Phi(\alpha_1)]
\end{aligned}
$$

Therefore, using the theorem above we get the following resuts:

$$
\begin{aligned}
\int_{a_1}^{a_2} y f(y)\, dy &= \mu [\Phi(\alpha_2) - \Phi(\alpha_1)] - \sigma [\phi(\alpha_2) - \phi(\alpha_1)] \\
\int_{a_1}^{a_2} y^2 f(y)\, dy &= \left[\sigma^2 + \mu^2\right] \cdot [\Phi(\alpha_2) - \Phi(\alpha_1)] - \sigma^2 [\alpha_2 \phi(\alpha_2) - \alpha_1 \phi(\alpha_1)] - 2\mu\sigma [\phi(\alpha_2) - \phi(\alpha_1)] \\
\int_{a_1}^{\infty} y f(y)\, dy &= \mu [1 - \Phi(\alpha_1)] + \sigma \phi(\alpha_1) \\
\int_{a_1}^{\infty} y^2 f(y)\, dy &= \left[\sigma^2 + \mu^2\right] \cdot [1 - \Phi(\alpha_1)] + \left[\sigma^2 \alpha_1 + 2\mu\sigma\right] \phi(\alpha_1)
\end{aligned}
$$

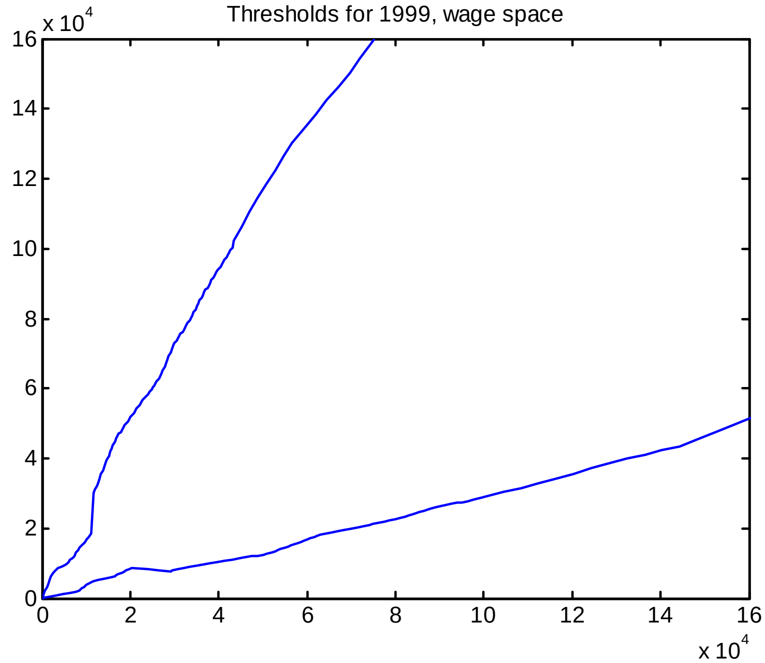
Theorem 3 *Let $Y \sim N[\mu, \sigma^2]$ with density $f(y)$. Then*

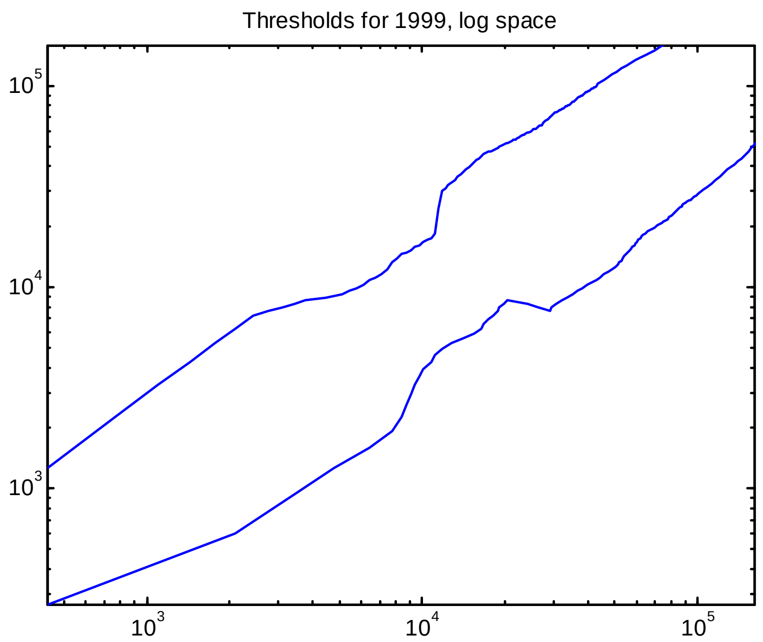
$$\int_{a_1}^{a_2} e^{ty} f(y)\, dy = \exp\left(\mu t + \sigma^2 t^2/2\right) \left[\Phi\left(\frac{a_2 - \mu}{\sigma} - \sigma t \right) - \Phi\left(\frac{a_1 - \mu}{\sigma} - \sigma t \right) \right]$$

[1] For a proof see Sam Cortum's notes at http://www.econ.umn.edu/~kortum/courses/fall02/lecture4k.pdf

3.3 3-choice model

In the 3-choice model the decision rules give rise to two thresholds:





Given the parameters of the model, we would like to efficiently compute the following moments:

Moment
1. $P(1M)$
2. $P(1F)$
3. $E[X\|1M \cup 2E]$
4. $Var[X\|1M \cup 2E]$
5. $E[Y\|1F \cup 2E]$
6. $Var[Y\|1F \cup 2E]$
7. $Cov[X, Y\|2E]$
8. $P(1M\|1M \cup 2E,\ X \leq a)$

It is possible that the taxes are such that for some low incomes the two thresholds intersect. In what follows we assume that in the relevant range we have $\forall x,\, h(x) > l(x)$.

1. $P(1M)$

$$\begin{aligned} P(1M) &= \int_{-\infty}^{\infty}\int_{-\infty}^{l(x)} f(x,y)\,dydx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{l(x)} f(y|x)\,f_X(x)\,dydx \\ &= \int_{-\infty}^{\infty} f_X(x)\int_{-\infty}^{l(x)} f(y|x)\,dydx \\ &= \int_{-\infty}^{\infty} F(l(x)|x)\,f_X(x)\,dx \end{aligned}$$

We can perform the standard transformation

$$z = \frac{x-\mu_X}{\sigma_X}, \text{ so that } x = z\sigma_X + \mu_X$$

and let

$$\alpha_1(x) = \frac{l(x) - \mu_{Y|X=x}}{\sigma_{Y|X=x}}$$

Thus,

$$\begin{aligned} P(1M) &= \int_{-\infty}^{\infty} \Phi(\alpha_1(x))\,f_X(x)\,dx \\ &= \int_{-\infty}^{\infty} \Phi(\alpha_1(x))\,\phi(z)\,dz \\ \text{where } x &= z\sigma_X + \mu_X \end{aligned}$$

2. $P\left(2E\right)$

$$\begin{aligned}
P\left(1F\right) &= \int_{-\infty}^{\infty}\int_{h(x)}^{\infty} f\left(x,y\right)dydx \\
&= \int_{-\infty}^{\infty}\int_{h(x)}^{\infty} f\left(y|x\right)f_X\left(x\right)dydx \\
&= \int_{-\infty}^{\infty} f_X\left(x\right)\int_{h(x)}^{\infty} f\left(y|x\right)dydx \\
&= \int_{-\infty}^{\infty} f_X\left(x\right)\left[1-F\left(h\left(x\right)|x\right)\right]dx
\end{aligned}$$

Let

$$\alpha_2\left(x\right)=\frac{h\left(x\right)-\mu_{Y|X=x}}{\sigma_{Y|X=x}}$$

$$\begin{aligned}
P\left(1F\right) &= \int_{-\infty}^{\infty}\left[1-\Phi\left(\alpha_2\left(x\right)\right)\right]f_X\left(x\right)dx \\
&= \int_{-\infty}^{\infty}\left[1-\Phi\left(\alpha_2\left(x\right)\right)\right]\phi\left(z\right)dz \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}$$

Thus, $P\left(2E\right)=1-P\left(1M\right)-P(1F)$.

3. $E\left[X|1M\cup 2E\right]$

$$\begin{aligned}
E\left[X|1M\right] &= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{h(x)} xf\left(x,y\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{h(x)} xf\left(y|x\right)f_X\left(x\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} xf_X\left(x\right)\int_{-\infty}^{h(x)} f\left(y|x\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} xf_X\left(x\right)F\left(h\left(x\right)|x\right)dx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} x\Phi\left(\alpha_2\left(x\right)\right)\phi\left(z\right)dz}{P\left(1M\right)+P\left(2E\right)} \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}$$

4. $Var\left[X|1M\cup 2E\right]$

$$\begin{aligned}
E\left[X^2|1M\cup 2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{h(x)} x^2 f\left(x,y\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{h(x)} x^2 f\left(y|x\right)f_X\left(x\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} x^2 f_X\left(x\right)\int_{-\infty}^{h(x)} f\left(y|x\right)dydx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} x^2 F\left(h\left(x\right)|x\right)f_X\left(x\right)dx}{P\left(1M\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} x^2 \Phi\left(\alpha_2\left(x\right)\right)\phi\left(z\right)dz}{P\left(1M\right)+P\left(2E\right)} \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}$$

Therfore

$$Var\left[X|1M\cup 2E\right]=E\left[X^2|1M\cup 2E\right]-E^2\left[X|1M\cup 2E\right]$$

5. $E\left[Y|1F\cup 2E\right]$

$$\begin{aligned}
E\left[Y|1F\cup 2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{\infty} yf\left(x,y\right)dydx}{P\left(1F\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{\infty} yf\left(y|x\right)f_X\left(x\right)dydx}{P\left(1F\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} f_X\left(x\right)\left[\int_{l(x)}^{\infty} yf\left(y|x\right)dy\right]dx}{P\left(1F\right)+P\left(2E\right)}
\end{aligned}$$

Let the integral in the brackets be $I_1\left(x\right)$. From theorem 2 we have

$$I_1\left(x\right)=\mu_{Y|X=x}\left[1-\Phi\left(\alpha_1\left(x\right)\right)\right]+\sigma_{Y|X=x}\phi\left(\alpha_1\left(x\right)\right)$$

Thus,

$$\begin{aligned}
E\left[Y|1F\cup 2E\right] &= \frac{\int_{-\infty}^{\infty} I_1\left(x\right)f_X\left(x\right)dx}{P\left(1F\right)+P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} I_1\left(x\right)\phi\left(z\right)dz}{P\left(1F\right)+P\left(2E\right)} \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}$$

6. $Var\left[Y|1F\cup 2E\right]$

$$\begin{aligned} E\left[Y^2|1F\cup 2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{\infty} y^2 f\left(x,y\right)dydx}{P\left(1F\right)+P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{\infty} y^2 f\left(y|x\right)f_X\left(x\right)dydx}{P\left(1F\right)+P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty} f_X\left(x\right)\left[\int_{l(x)}^{\infty} y^2 f\left(y|x\right)dy\right]dx}{P\left(1F\right)+P\left(2E\right)} \end{aligned}$$

Let the integral in the brackets be $I_2\left(x\right)$. Then, from theorem 2 we have

$$I_2\left(x\right)=\left[\sigma_{Y|X=x}^2+\mu_{Y|X=x}^2\right]\cdot\left[1-\Phi\left(\alpha_1\left(x\right)\right)\right]+\left[\sigma_{Y|X=x}^2\alpha_1\left(x\right)+2\mu_{Y|X=x}\sigma_{Y|X=x}\right]\phi\left(\alpha_1\left(x\right)\right)$$

Thus,

$$\begin{aligned} E\left[Y^2|2E\right] &= \frac{\int_{-\infty}^{\infty} I_2\left(x\right)f_X\left(x\right)dx}{P\left(1F\right)+P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty} I_2\left(x\right)\phi\left(z\right)dx}{P\left(1F\right)+P\left(2E\right)} \\ \text{where } x &= z\sigma_X+\mu_X \end{aligned}$$

And

$$Var\left[Y|1F\cup 2E\right]=E\left[Y^2|1F\cup 2E\right]-E^2\left[Y|1F\cup 2E\right]$$

7. $Cov\left[X,Y|2E\right]$

$Cov\left[X,Y|2E\right]=E\left[XY|2E\right]-E\left[X|2E\right]E\left[Y|2E\right]$, where

$$\begin{aligned} E\left[XY|2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} xyf\left(x,y\right)dydx}{P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} xyf\left(y|x\right)f_X\left(x\right)dydx}{P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty} xf_X\left(x\right)\left[\int_{l(x)}^{h(x)} yf\left(y|x\right)dy\right]dx}{P\left(2E\right)} \end{aligned}$$

Let the integral in the brackets be $I_3\left(x\right)$. Then, from theorem 2 we have

$$I_3\left(x\right)=\mu_{Y|X=x}^2\cdot\left[\Phi\left(\alpha_2\left(x\right)\right)-\Phi\left(\alpha_1\left(x\right)\right)\right]-\sigma_{Y|X=x}\left[\phi\left(\alpha_2\left(x\right)\right)-\phi\left(\alpha_1\left(x\right)\right)\right]$$

Thus,

$$\begin{aligned} E\left[XY|2E\right] &= \frac{\int_{-\infty}^{\infty} xI_3\left(x\right)f_X\left(x\right)dx}{P\left(2E\right)} \\ &= \frac{\int_{-\infty}^{\infty} xI_3\left(x\right)\phi\left(z\right)dz}{P\left(2E\right)} \\ \text{where } x &= z\sigma_X+\mu_X \end{aligned}$$

Now,

$$
\begin{aligned}
E\left[X|2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} xf\left(x,y\right)dydx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} xf\left(y|x\right)f_X\left(x\right)dydx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} xf_X\left(x\right)\int_{l(x)}^{h(x)} f\left(y|x\right)dydx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} xf_X\left(x\right)\left[F\left(h\left(x\right)|x\right)-F\left(l\left(x\right)|x\right)\right]dx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} x\left[\Phi\left(\alpha_2\left(x\right)\right)-\Phi\left(\alpha_1\left(x\right)\right)\right]\phi\left(z\right)dz}{P\left(2E\right)} \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}
$$

And finally,

$$
\begin{aligned}
E\left[Y|2E\right] &= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} yf\left(x,y\right)dydx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty}\int_{l(x)}^{h(x)} yf\left(y|x\right)f_X\left(x\right)dydx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} f_X\left(x\right)\left[\int_{l(x)}^{h(x)} yf\left(y|x\right)dy\right]dx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} f_X\left(x\right)I_3\left(x\right)dx}{P\left(2E\right)} \\
&= \frac{\int_{-\infty}^{\infty} I_3\left(x\right)\phi\left(z\right)dz}{P\left(2E\right)} \\
\text{where } x &= z\sigma_X+\mu_X
\end{aligned}
$$

8. $P(1M|1M \cup 2E,\ X \leq a)$

$$\begin{aligned}
P(1M|1M \cup 2E,\ X \leq a) &= \frac{P(1M,\ 1M \cup 2E,\ X \leq a)}{P(1M \cup 2E,\ X \leq a)} \\
&= \frac{P(Y \leq l(x),\ X \leq a)}{P(Y \leq h(x),\ X \leq a)} \\
&= \frac{\int_{-\infty}^{a} \int_{-\infty}^{l(x)} f(x,y)\, dydx}{\int_{-\infty}^{a} \int_{-\infty}^{h(x)} f(x,y)\, dydx} \\
&= \frac{\int_{-\infty}^{a} \int_{-\infty}^{l(x)} f(y|x)\, f_X(x)\, dydx}{\int_{-\infty}^{a} \int_{-\infty}^{h(x)} f(y|x)\, f_X(x)\, dydx} \\
&= \frac{\int_{-\infty}^{a} f_X(x) \int_{-\infty}^{l(x)} f(y|x)\, dydx}{\int_{-\infty}^{a} f_X(x) \int_{-\infty}^{h(x)} f(y|x)\, dydx} \\
&= \frac{\int_{-\infty}^{a} F(l(x)|x)\, f_X(x)\, dx}{\int_{-\infty}^{a} F(h(x)|x)\, f_X(x)\, dx} \\
\text{where } x &= z\sigma_X + \mu_X
\end{aligned}$$

In particular, we can choose $a = \mu_X$.

3.4 Other integrals of interest

1. Wife's participation by interval of husband's income: $P(2E|1M \cup 2E,\ a \leq X \leq b)$. We compute the probability of the complement.

$$\begin{aligned}
P(1M|1M \cup 2E,\ a \leq X \leq b) &= \frac{P(1M,\ 1M \cup 2E,\ a \leq X \leq b)}{P(1M \cup 2E,\ a \leq X \leq b)} \\
&= \frac{P(1M,\ a \leq X \leq b)}{P(1M \cup 2E,\ a \leq X \leq b)} \\
&= \frac{P(Y \leq l(x),\ a \leq X \leq b)}{P(Y \leq h(x),\ a \leq X \leq b)} \\
&= \frac{\int_{a}^{b} \int_{-\infty}^{l(x)} f(x,y)\, dydx}{\int_{a}^{b} \int_{-\infty}^{h(x)} f(x,y)\, dydx} \\
&= \frac{\int_{a}^{b} \int_{-\infty}^{l(x)} f(y|x)\, f_X(x)\, dydx}{\int_{a}^{b} \int_{-\infty}^{h(x)} f(y|x)\, f_X(x)\, dydx} \\
&= \frac{\int_{a}^{b} f_X(x) \left[\int_{-\infty}^{l(x)} f(y|x)\, dy\right] dx}{\int_{a}^{b} f_X(x) \left[\int_{-\infty}^{h(x)} f(y|x)\, dy\right] dx} \\
&= \frac{\int_{a}^{b} f_X(x)\, [F(l(x)|x)]\, dx}{\int_{a}^{b} f_X(x)\, [F(h(x)|x)]\, dx}
\end{aligned}$$

We would like to copute the above probability for intervals of X defined by the nodes $-\infty, z_1, z_2, ..., z_n, \infty$.

2. Suppose that we want to examine closely the impact of the earned income tax credit on the low income families. In particular, we are interested in computing $P(1M|X \le a,\ Y \le b)$.

$$\begin{aligned}
P(1M|X \le a,\ Y < b) &= \frac{P(1M,\ X \le a,\ Y \le b)}{P(X \le a,\ Y \le b)} \\
&= \frac{P(Y \le l(x),\ X \le a,\ Y \le b)}{P(X \le a,\ Y \le b)} \\
&= \frac{P(Y \le \min\{l(x), b\},\ X \le a)}{P(X \le a,\ Y \le b)} \\
&= \frac{\int_{-\infty}^{a} \int_{-\infty}^{l(x)\wedge b} f(x,y)\, dydx}{P(X \le a,\ Y < b)} \\
&= \frac{\int_{-\infty}^{a} \int_{-\infty}^{l(x)\wedge b} f(y|x) f_X(x)\, dydx}{\int_{-\infty}^{a} \int_{-\infty}^{b} f(y|x) f_X(x)\, dydx} \\
&= \frac{\int_{-\infty}^{a} f_X(x) \int_{-\infty}^{l(x)\wedge b} f(y|x)\, dydx}{\int_{-\infty}^{a} f_X(x) \int_{-\infty}^{b} f(y|x)\, dydx} \\
&= \frac{\int_{-\infty}^{a} F(l(x) \wedge b|x) f_X(x)\, dx}{\int_{-\infty}^{a} F(b|x) f_X(x)\, dx}
\end{aligned}$$

It may be easier to just compute this probability by simulating a large sample.

4 Simulating random draws from normal distribution

Often we need to simulate a large sample from multivariate normal (or LogNormal) distribution. Sometimes we need to compute an approximation to a moment just once and we don't want to write a code for numerical integration. Or, sometimes it is very difficult to compute the integral numerically, because of high dementionality of the integrand. Yet another use of random draws is forchecking the numerical computation.

Any statistical or mathematical software, has random number generator from the univariate standard normal distribution. So we can obtain a vector of uncorrelated standard normal r.v.'s

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

where $Z_i \sim N(0,1)$. Now, suppose that we need a random draw from general normal distribution, i.e.,

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Then we let

$$X = \boldsymbol{\mu} + \mathbf{P}Z$$

where $\mathbf{P}$ is the lower Cholesky decomposition of $\boldsymbol{\Sigma}$, so that $\mathbf{PP}' = \boldsymbol{\Sigma}$. To verify that $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, recall that X is a linear function of Z, so it has to be normal. The only thing left to do is to compute the mean and variance

$$\begin{aligned} E(X) &= \boldsymbol{\mu} \\ Var(X) &= \mathbf{P}Var(Z)\mathbf{P}' = \mathbf{PIP}' = \boldsymbol{\Sigma} \end{aligned}$$

5 Relating moments of the Normal and LogNormal

Suppose that $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and we define $Y = \exp(X)$, so that Y has LogNormal distribution. In other words,

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \exp(X_1) \\ \vdots \\ \exp(X_n) \end{bmatrix}$$

We need to find the mean and the covariance matrix of Y. Recall that the moment generating function of a random variable is defined as follows.

$$\psi(\mathbf{t}) = E(\exp(\mathbf{t}X)) = E(\exp(t_1 X_1 + ... + t_n X_n))$$

Let denote the mean vector and the covariance matrix of Y by $\mathbf{m}$ and $\mathbf{S}$ respectively. To find the mean vector observe that

$$E(Y_i) = E(\exp(X_i))$$

Thus, we simply let $t_i = 1$ and $t_j = 0\ \forall j \neq i$ and evaluate the m.g.f. at this $\mathbf{t}$.

To find the second moments observe that

$$E\left(Y_i^2\right) = E\left(e^{X_i} e^{X_i}\right) = E(\exp(2X_i))$$

Thus we let $t_i = 2$ and $t_j = 0\ \forall j \neq i$ and evaluate the m.g.f. at this $\mathbf{t}$. Then the variance of the i^{th} component is obtained by

$$var(Y_i) = E\left(Y_i^2\right) - E^2(Y_i)$$

Finally, the covariance is obtained by

$$\begin{aligned} cov(Y_i, Y_j) &= E(Y_i Y_j) - E(Y_i)E(Y_j) \\ &= E\left(e^{X_i} e^{X_j}\right) - E\left(e^{X_i}\right)E\left(e^{X_j}\right) \\ &= E(\exp(X_i + X_j)) - E(\exp(X_i))E(\exp(X_j)) \end{aligned}$$

To find the first term on the right we set $t_i = t_j = 1$, and the other coordinates in the vector $\mathbf{t}$ are set to 0.

Recall that the moment generating function of multivariate normal is

$$\psi(\mathbf{t}) = \exp\left(\mathbf{t}'\boldsymbol{\mu}+\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right)$$

Therefore, finding moments of the LogNormal distribution is a simple task of evaluating the m.g.f. of the normal distribution at different vectors $\mathbf{t}$, as described above.

5.1 A two variables example

Let $X = [X_1, X_2] \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \ \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

Let $Y = [\exp(X_1), \exp(X_2)] \sim LN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We want to find the mean vector, $\mathbf{m}$, and the covariance matrix, $\mathbf{S}$, of Y, explicitly written as

$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \ \mathbf{S} = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix}$$

Following the discussion in the previous section we get:
(m_1)

$$\begin{aligned} m_1 &= \psi\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \\ &= \exp\left(\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}+\frac{1}{2}\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \\ &= \exp\left(\mu_1 + \frac{1}{2}\sigma_{11}\right) \end{aligned}$$

(m_2)

$$\begin{aligned} m_2 &= \psi\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= \exp\left(\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}+\frac{1}{2}\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= \exp\left(\mu_2 + \frac{1}{2}\sigma_{22}\right) \end{aligned}$$

(s_1^2)

$$\begin{aligned} E\left(Y_1^2\right) &= \psi\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) \\ &= \exp\left(\begin{bmatrix} 2 & 0 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}+\frac{1}{2}\begin{bmatrix} 2 & 0 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) \\ &= \exp(2\mu_1 + 2\sigma_{11}) \\ var(Y_1) &= \exp(2\mu_1 + 2\sigma_{11}) - \exp(2\mu_1 + \sigma_{11}) = \exp(2\mu_1 + \sigma_{11})(\exp(\sigma_{11}) - 1) \end{aligned}$$

(s_2^2)

$$
\begin{aligned}
E\left(Y_2^2\right) &= \psi\left(\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right) \\
&= \exp\left(\begin{bmatrix} 0 & 2 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 & 2 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right) \\
&= \exp\left(2\mu_2 + 2\sigma_{22}\right) \\
var\left(Y_2\right) &= \exp\left(2\mu_2 + 2\sigma_{22}\right) - \exp\left(2\mu_2 + \sigma_{22}\right) = \exp\left(2\mu_2 + \sigma_{22}\right)\left(\exp\left(\sigma_{22}\right) - 1\right)
\end{aligned}
$$

(s_{12})

$$
\begin{aligned}
E\left(Y_1 Y_2\right) &= \psi\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \\
&= \exp\left(\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \\
&= \exp\left(\mu_1 + \mu_2 + \sigma_{12} + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right) \\
cov\left(Y_1, Y_2\right) &= \exp\left(\mu_1 + \mu_2 + \sigma_{12} + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right) - \exp\left(\mu_1 + \frac{1}{2}\sigma_{11}\right)\exp\left(\mu_2 + \frac{1}{2}\sigma_{22}\right) \\
&= \exp\left(\mu_1 + \mu_2 + \sigma_{12} + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right) - \exp\left(\mu_1 + \mu_2 + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right) \\
&= \exp\left(\mu_1 + \mu_2 + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right)\left(\exp\left(\sigma_{12}\right) - 1\right)
\end{aligned}
$$

Summary of the parameters:

$$
\begin{aligned}
m_1 &= \exp\left(\mu_1 + \frac{1}{2}\sigma_{11}\right) \\
m_2 &= \exp\left(\mu_2 + \frac{1}{2}\sigma_{22}\right) \\
s_1^2 &= \exp\left(2\mu_1 + \sigma_{11}\right)\left(\exp\left(\sigma_{11}\right) - 1\right) \\
s_2^2 &= \exp\left(2\mu_2 + \sigma_{22}\right)\left(\exp\left(\sigma_{22}\right) - 1\right) \\
s_{12} &= \exp\left(\mu_1 + \mu_2 + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22}\right)\left(\exp\left(\sigma_{12}\right) - 1\right)
\end{aligned}
$$

5.1.1 Solving for μ_i and σ_{ij}

We now solve analytially for the parameters of the underlying normal distribution given the parameters of the LogNormal.

$$\begin{aligned}
\log m_1 &= \mu_1 + \frac{1}{2}\sigma_{11} \\
\log m_2 &= \mu_2 + \frac{1}{2}\sigma_{22} \\
\log s_1^2 &= 2\mu_1 + \sigma_{11} + \log(\exp(\sigma_{11}) - 1) \\
\log s_2^2 &= 2\mu_2 + \sigma_{22} + \log(\exp(\sigma_{22}) - 1) \\
\log s_{12} &= \mu_1 + \mu_2 + \frac{1}{2}\sigma_{11} + \frac{1}{2}\sigma_{22} + \log(\exp(\sigma_{12}) - 1)
\end{aligned}$$

(σ_1^2)

$$\begin{aligned}
\log s_{11} &= 2\mu_1 + \sigma_{11} + \log(\exp(\sigma_{11}) - 1) \\
\log s_{11} &= 2\log m_1 + \log(\exp(\sigma_{11}) - 1) \\
\log\left(\frac{s_{11}}{m_1^2}\right) &= \log(\exp(\sigma_{11}) - 1) \\
\frac{s_{11}}{m_1^2} &= \exp(\sigma_{11}) - 1 \\
\sigma_{11} &= \log\left(1 + \frac{s_{11}}{m_1^2}\right)
\end{aligned}$$

(σ_2^2)

$$\sigma_{22} = \log\left(1 + \frac{s_{22}}{m_2^2}\right)$$

(μ_1)

$$\begin{aligned}
\log m_1 &= \mu_1 + \frac{1}{2}\sigma_{11} \\
\mu_1 &= \log m_1 - \frac{1}{2}\log\left(1 + \frac{s_{11}}{m_1^2}\right) \\
\mu_1 &= \log m_1 - \log\left(\frac{\sqrt{m_1^2 + s_{11}}}{m_1}\right) \\
\mu_1 &= \log\left(\frac{m_1^2}{\sqrt{m_1^2 + s_{11}}}\right)
\end{aligned}$$

(μ_2)

$$\mu_2 = \log\left(\frac{m_2^2}{\sqrt{m_2^2 + s_{22}}}\right)$$

(σ_{12})

$$
\begin{aligned}
\log m_1 &= \mu_1 + \frac{1}{2}\sigma_{11} \\
\log m_2 &= \mu_2 + \frac{1}{2}\sigma_{22} \\
\log m_1 + \log m_2 - \frac{1}{2}\sigma_{11} - \frac{1}{2}\sigma_{22} &= \mu_1 + \mu_2 \\
\log s_{12} &= \log m_1 + \log m_2 + \log\left(\exp\left(\sigma_{12}\right) - 1\right) \\
\log s_{12} &= \log\left(m_1 m_2 \left(\exp\left(\sigma_{12}\right) - 1\right)\right) \\
s_{12} &= m_1 m_2 \left(\exp\left(\sigma_{12}\right) - 1\right) \\
\sigma_{12} &= \log\left(1 + \frac{s_{12}}{m_1 m_2}\right)
\end{aligned}
$$

Summary

$$
\begin{aligned}
\mu_1 &= \log\left(\frac{m_1^2}{\sqrt{m_1^2 + s_{11}}}\right) \\
\mu_2 &= \log\left(\frac{m_2^2}{\sqrt{m_2^2 + s_{22}}}\right) \\
\sigma_1^2 &= \log\left(1 + \frac{s_{11}}{m_1^2}\right) \\
\sigma_2^2 &= \log\left(1 + \frac{s_{22}}{m_2^2}\right) \\
\sigma_{12} &= \log\left(1 + \frac{s_{12}}{m_1 m_2}\right)
\end{aligned}
$$

5.2 General formulas

Given the Normal random variable $X \sim N\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$, we want to find the mean and covariance matrix of the LogNormal $Y = \exp\left(X\right)$. Let the mean vector and covariance matrix of Y be $\left(\mathbf{m}, \mathbf{S}\right)$

$$
\begin{aligned}
E\left(Y_i\right) &= \exp\left(\mu_i + \frac{\sigma_{ii}}{2}\right) \\
Var\left(Y_i\right) &= \exp\left(2\mu_i + \sigma_{ii}\right)\left(\exp\left(\sigma_{ii}\right) - 1\right) \\
Cov\left(Y_i, Y_j\right) &= \left(\exp\left(\sigma_{ij}\right) - 1\right)\exp\left(\mu_i + \mu_j + \frac{\sigma_{ii} + \sigma_{jj}}{2}\right) \\
Cor\left(Y_i, Y_j\right) &= \frac{\exp\left(\sigma_{ij}\right) - 1}{\sqrt{\left(\exp\left(\sigma_{ii}\right) - 1\right)\left(\exp\left(\sigma_{jj}\right) - 1\right)}}
\end{aligned}
$$

Given the mean vector and the covariance matrix of the LogNormal distribution, we obtain the mean and covaraince matrix of the Normal random variable as follows

$$\begin{aligned}\mu_i &= E(X_i) = \ln\left(\frac{m_i^2}{\sqrt{m_i^2 + s_{ii}}}\right) \\ \sigma_{ij} &= Cov(X_i, X_j) = \ln\left(1 + \frac{s_{ij}}{m_i m_j}\right)\end{aligned}$$

6 Maximum Likelihood Estimation

In the benchmark version of the model, the households choose

$$\max\{V_{2E}(w_m, w_f), V_{1M}(w_m, w_f), V_{1F}(w_m, w_f)\},$$

where

$$\begin{gathered}V_{2E}(w_m, w_f) = \\ \max_{c_m, c_f}\{\lambda[\alpha\log(c_m) + (1-\alpha)\log(1-l_m)] + (1-\lambda)[\alpha\log(c_f) + (1-\alpha)\log(1-l_f)]\} \\ \text{s.t.} \quad (1+\tau)(c_m + c_f) = w_m + w_f - T(w_m, w_f),\end{gathered}$$

$$\begin{gathered}V_{1M}(w_m, w_f) = \\ \max_{c_m, c_f}\{\lambda[\alpha\log(c_m) + (1-\alpha)\log(1-l_m)] + (1-\lambda)\alpha\log(c_f)\} \\ \text{s.t.} \quad (1+\tau)(c_m + c_f) = w_m - T(w_m, 0),\end{gathered}$$

$$\begin{gathered}V_{1F}(w_m, w_f) = \\ \max_{c_m, c_f}\{\lambda\alpha\log(c_m) + (1-\lambda)[\alpha\log(c_f) + (1-\alpha)\log(1-l_f)]\} \\ \text{s.t.} \quad (1+\tau)(c_m + c_f) = w_f - T(0, w_f).\end{gathered}$$

A word about notation. There is a difference between $V_{2E}(w_m, w_f), V_{1M}(w_m, w_f), V_{1F}(w_m, w_f)$ and V_{2E}, V_{1M}, V_{1F}. The former are deterministic values and describe the value for *given wages*. However, when we write V_{2E}, V_{1M}, V_{1F} we are referring to random variables since the wages are random. In what follows, it will be convenient to write the maximum value functions as functions of the log of wages, i.e., $V_{2E}(x, y), V_{1M}(x, y), V_{1F}(x, y)$. Thus, In the original model the household chooses $\max\{V_{2E}(x, y), V_{1M}(x, y), V_{1F}(x, y)\}$.

6.1 Empirical version of the model

Now suppose that we modify this slightly so that the household chooses

$$\begin{aligned}&\max\{V_{2E}(x, y) + \varepsilon_1, V_{1M}(x, y) + \varepsilon_2, V_{1F}(x, y) + \varepsilon_3\} \\ \equiv\ &\max\left\{\tilde{V}_{2E}(x, y), \tilde{V}_{1M}(x, y), \tilde{V}_{1F}(x, y)\right\}\end{aligned}$$

where ε_i are $i.i.d.$, $\sim N(0,\sigma)$. This implies that $\tilde{V}_{2E}(x,y), \tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y)$ are $i.i.d.$random variables, with normal distribution

$$\begin{aligned}\tilde{V}_{2E}(x,y) &\sim N(V_{2E},\sigma)\\ \tilde{V}_{1M}(x,y) &\sim N(V_{1M},\sigma)\\ \tilde{V}_{1F}(x,y) &\sim N(V_{1F},\sigma)\end{aligned}$$

We call this version "stochastic", and the interpretation of the above modification is as follows. There are factors other than the spouses' incomes that affect the labor force participation decision. These factors are not in the model, but we realize that they exist. The means of $\tilde{V}_{2E}, \tilde{V}_{1M}, \tilde{V}_{1F}$ are V_{2E}, V_{1M}, V_{1F}, so that on average couples will tend to make choices according to the thresholds of the deterministic version. We can still draw the 3 regions that we called $2E, 1M, 1F$ and use the picture for motivating. The advantage however is that in the stochastic version, it is possible that even if (w_m, w_f) falls into the $1M$ region, the couple will be a 2-earner couple. Also, there is a positive probability (likelihood) to obtain the data from this model. The cost of this extension is an additional parameter to calibrate: σ.

6.2 Likelihood function

Let the log of the potential wage of couple i be $(\tilde{x}^i, \tilde{y}^i)$ and let the joint density be $f(\tilde{x}^i, \tilde{y}^i)$. We can have three types of observations, as described below.

$$\left(x^i, y^i\right) = \begin{cases} (\tilde{x}^i, \tilde{y}^i) & \text{if the couple chose to be } 2E \\ (\tilde{x}^i, \cdot) & \text{if the couple chose to be } 1M \\ (\cdot, \tilde{y}^i) & \text{if the couple chose to be } 1F \end{cases}$$

Conditional on observing $(\tilde{x}^i, \tilde{y}^i)$, the density is $f(x^i, y^i|2E)$, which we derive using the Bayes rule[2] as follows.

$$\begin{aligned} f\left(x^i, y^i|2E\right) P(2E) &= P\left(2E|x^i, y^i\right) f\left(x^i, y^i\right) \\ f\left(x^i, y^i|2E\right) &= \frac{P(2E|x^i, y^i) f(x^i, y^i)}{P(2E)} \end{aligned}$$

Conditional on observing $(\tilde{x}^i, \cdot)$, the density is

$$\begin{aligned} f\left(x^i|1M\right) P(1M) &= P\left(1M|x^i\right) f_X(x) \\ f\left(x^i|1M\right) &= \frac{P(1M|x^i) f_X(x)}{P(1M)} \end{aligned}$$

And conditional on observing $(\cdot, \tilde{y}^i)$, the density is

$$\begin{aligned} f\left(y^i|1F\right) P(1F) &= P\left(1F|y^i\right) f_Y\left(y^i\right) \\ f\left(y^i|1F\right) &= \frac{P(1F|y^i) f_Y(y^i)}{P(1F)} \end{aligned}$$

[2] We prove this in the appendix.

However, the contribution to the likelihood function made by the observations of the first type, $(\tilde{x}^i, \tilde{y}^i)$, is not $f(x^i, y^i|2E)$, because these observations occur only with probability $P(2E)$. The contribution to the likelihood is therefore $f(x^i, y^i|2E)\,P(2E)$. Similarly, the contribution to the likelihood by the observations of the second kind is $f(x^i|1M)\,P(1M)$ and the contribution observations of the third type is $f(y^i|1F)\,P(1F)$. Thus, the log likelihood function is

$$l = \sum_{i\in 2E} \log\left(P\left(2E|x^i, y^i\right) f\left(x^i, y^i\right)\right) + \sum_{i\in 1M} \log\left(P\left(1M|x^i\right) f_X\left(x^i\right)\right) + \sum_{i\in 1F} \log\left(P\left(1F|y^i\right) f_Y\left(y^i\right)\right)$$

Intuition. Suppose that we write the following (wrong) log likelihood function

$$l = \sum_{i\in 2E} \log\left(f\left(x^i, y^i\right)\right) + \sum_{i\in 1M} \log\left(f_X\left(x^i\right)\right) + \sum_{i\in 1F} \log\left(f_Y\left(y^i\right)\right)$$

At first site this might look like a reasonable choice, and it is much easier to maximize than the correct one. Notice however that in this log likelihood the economic model does not play any role. This is the log likelihood that we would write if the observations were arbitrary censored and we had no clue why in some observations we do observe both x and y, and sometimes we don't. However, since we do have a model, then when we observe both x and y, our intuition tells us us that there is smaller chance that x is very big relative to y or vice versa. Similarly, when we observe only x, then our intuition tells us that most likely x is relatively big. In the correct likelihood we see that our intuition is implemented. The densities are weighted by some probabilities. So for example take the second term in the correct log likelihood: $P(1M|x^i)\,f_X(x^i)$. Suppose that we observe a small x. Then the probability that this couple is $1M$ should be small, and the term $P(1M|x^i)$ takes care of that exactly. This probability is small if x^i is small and it is high if x is high so the term $P(1M|x^i)$ puts weights on the density $f_X(x^i)$ depending on the magnitude of x^i and in accordance with the model.

The rest of these notes is a detailed explanation of how to implement the technique, i.e., how to compute the log likelihood function in practice.

6.2.1 Computing the log likelihood, theory

1. $P(2E|x, y)$

$$\begin{aligned}
P(2E|x,y) &= P\left(\max\left\{\tilde{V}_{2E}(x,y), \tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y)\right\} = \tilde{V}_{2E}(2E)\,|x,y\right) \\
&= P\left(\tilde{V}_{1M}(x,y) \le \tilde{V}_{2E}(x,y), \tilde{V}_{1F}(x,y) \le \tilde{V}_{2E}(x,y)\,|x,y\right) \\
&= \int P\left(\tilde{V}_{1M}(x,y) \le s, \tilde{V}_{1F}(x,y) \le s|x,y\right) f_{2E}(s)\,ds \\
&= \int \Phi\left(\frac{s - V_{1M}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{s - V_{1F}(x,y)}{\sigma}\right) f_{2E}(s)\,ds \\
&= \int \Phi\left(\frac{z\sigma + V_{2E}(x,y) - V_{1M}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{z\sigma + V_{2E}(x,y) - V_{1F}(x,y)}{\sigma}\right) \phi(z)\,dz \\
&= \int \Phi\left(z + \frac{V_{2E}(x,y) - V_{1M}(x,y)}{\sigma}\right) \cdot \Phi\left(z + \frac{V_{2E}(x,y) - V_{1F}(x,y)}{\sigma}\right) \phi(z)\,dz
\end{aligned}$$

The key step is the third row, and it might need more elucidation. We start by computing $P\left(\tilde{V}_{1M}(x,y) \leq s, \tilde{V}_{1F}(x,y) \leq s|x,y\right)$ for given realization s of $\tilde{V}_{2E}(x,y)$. Given (x,y) we know that the two random variables $\left(\tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y)\right)$ are independent normal with

$$\begin{aligned} \tilde{V}_{1M}(x,y) &\sim N\left(V_{1M}(x,y), \sigma\right) \\ \tilde{V}_{1F}(x,y) &\sim N\left(V_{1F}(x,y), \sigma\right) \end{aligned}$$

Therefore,

$$\begin{aligned} P\left(\tilde{V}_{1M}(x,y) \leq s, \tilde{V}_{1F}(x,y) \leq s|x,y\right) &= P\left(\tilde{V}_{1M}(x,y) \leq s|x,y\right) \cdot P\left(\tilde{V}_{1F}(x,y) \leq s|x,y\right) \\ &= \Phi\left(\frac{s - V_{1M}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{s - V_{1F}(x,y)}{\sigma}\right) \end{aligned}$$

Now, since $\tilde{V}_{2E}(x,y)$ is in fact a random variable, we need to average the above expression over all its realizations, which is achieved by

$$\int \Phi\left(\frac{s - V_{1M}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{s - V_{1F}(x,y)}{\sigma}\right) f_{2E}(s)\, ds$$

where $f_{2E}(s)$ is the density of $\tilde{V}_{2E}(x,y) \sim N\left(V_{2E}(x,y), \sigma\right)$. Instead of integrating over s we let

$$Z = \frac{\tilde{V}_{2E}(x,y) - V_{2E}(x,y)}{\sigma} \sim N(0,1)$$

and thus $s = z\sigma + V_{2E}(x,y)$. Substituting this in the above integral and using the density of the standard normal, gives the final result.

2. $P(1M|x)$

We start by computing the probability of being a $1M$ couple, pretending that we observe both wages. That is, we first compute $P(1M|x,y)$.

$$\begin{aligned} P(1M|x,y) &= P\left(\max\left\{\tilde{V}_{2E}(x,y), \tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y)\right\} = \tilde{V}_{1M}(x,y)\,|x,y\right) \\ &= P\left(\tilde{V}_{2E}(x,y) \leq \tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y) \leq \tilde{V}_{1M}(x,y)\,|x,y\right) \\ &= \int P\left(\tilde{V}_{2E}(x,y) \leq s, \tilde{V}_{1F}(x,y) \leq s|x,y\right) f_{1M}(s)\, ds \\ &= \int \Phi\left(\frac{s - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{s - V_{1F}(x,y)}{\sigma}\right) f_{1M}(s)\, ds \\ &= \int \Phi\left(\frac{z\sigma + V_{1M}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{z\sigma + V_{1M}(x,y) - V_{1F}(x,y)}{\sigma}\right) \phi(z)\, dz \\ &= \int \Phi\left(z + \frac{V_{1M}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(z + \frac{V_{1M}(x,y) - V_{1F}(x,y)}{\sigma}\right) \phi(z)\, dz \end{aligned}$$

Now we need to average the above probability over all the values of y. The question is what weight should we put on every value of y? At first we might think that the unconditional density $f_Y(y)$ is the answer. But remember that in the second type of observations we do observe x, which contains some information about y, if they are correlated. Therefore, the correct weighting function is the conditional density of y given x, that is $f(y|x)$. Thus,

$$P(1M|x) = \int P(1M|x,y)\, f(y|x)\, dy$$

3. $P(1F|y)$

We start by computing the probability of being a $1F$ couple, pretending that we observe both wages. That is, we first compute $P(1F|x,y)$.

$$\begin{aligned}
P(1F|x,y) &= P\left(\max\left\{\tilde{V}_{2E}(x,y), \tilde{V}_{1M}(x,y), \tilde{V}_{1F}(x,y)\right\} = \tilde{V}_{1F}(x,y)\,|x,y\right) \\
&= P\left(\tilde{V}_{2E}(x,y) \leq \tilde{V}_{1F}(x,y), \tilde{V}_{1M}(x,y) \leq \tilde{V}_{1F}(x,y)\,|x,y\right) \\
&= \int P\left(\tilde{V}_{2E}(x,y) \leq s, \tilde{V}_{1M}(x,y) \leq s|x,y\right) f_{1F}(s)\, ds \\
&= \int \Phi\left(\frac{s - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{s - V_{1M}(x,y)}{\sigma}\right) f_{1F}(s)\, ds \\
&= \int \Phi\left(\frac{z\sigma + V_{1F}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(\frac{z\sigma + V_{1F}(x,y) - V_{1M}(x,y)}{\sigma}\right) \phi(z)\, dz \\
&= \int \Phi\left(z + \frac{V_{1F}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(z + \frac{V_{1F}(x,y) - V_{1M}(x,y)}{\sigma}\right) \phi(z)\, dz
\end{aligned}$$

and

$$P(1F|y) = \int P(1F|x,y)\, f(x|y)\, dx$$

Notice that $P(2E|x,y) + P(1M|x,y) + P(1F|x,y) = 1\ \forall(x,y)$, so in practice we can compute only two of those probabilities. But for checking that we don't have errors, we need to make sure that they sum up to 1.

6.2.2 Computing the log likelihood, practice

Suppose that we start with the first type of observations; those in which we observe both x and y. For each observation we compute $V_{2E}(x,y)$, $V_{1M}(x,y)$, and $V_{1F}(x,y)$. It is important for this step to have a continuous tax function, $T(x,y)$, which gives the tax liability for any given (x,y). Remember that TAXSIM was used to obtain tax returns for some grid points only. The tax function can be obtained by using (2-dimentional) interpolation over that grid. The output after this step is as follows

Obs. #	x	y	$V_{2E}(x,y)$	$V_{1M}(x,y)$	$V_{1F}(x,y)$
1	x^1	y^1	$V_{2E}(x^1,y^1)$	$V_{1M}(x^1,y^1)$	$V_{1F}(x^1,y^1)$
2	x^2	y^2	$V_{2E}(x^2,y^2)$	$V_{1M}(x^2,y^2)$	$V_{1F}(x^2,y^2)$
3	x^3	y^3	$V_{2E}(x^3,y^3)$	$V_{1M}(x^3,y^3)$	$V_{1F}(x^3,y^3)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Next we need to compute for each observation

$$P\left(2E|x,y\right)=\int\Phi\left(z+\frac{V_{2E}\left(x,y\right)-V_{1M}\left(x,y\right)}{\sigma}\right)\cdot\Phi\left(z+\frac{V_{2E}\left(x,y\right)-V_{1F}\left(x,y\right)}{\sigma}\right)\phi\left(z\right)dz$$

Note that

$$\Phi\left(z+\frac{V_{2E}\left(x,y\right)-V_{1M}\left(x,y\right)}{\sigma}\right)\cdot\Phi\left(z+\frac{V_{2E}\left(x,y\right)-V_{1F}\left(x,y\right)}{\sigma}\right)=P(2E|x,y,z)$$

and therefore

$$P\left(2E|x,y\right)=\int P(2E|x,y,z)\phi\left(z\right)dz$$

Notice that the integrand is a smooth function of z, regardless of how non-smooth the tax function is. This is because for each observation, the values of $V_{2E}\left(x,y\right)$, $V_{1M}\left(x,y\right)$, and $V_{1F}\left(x,y\right)$ are constant numbers, so the above integrand is just $\Phi\left(z+a\right)\cdot\Phi\left(z+b\right)\phi\left(z\right)$, and the functions Φ and ϕ are smooth. Therefore, we can use the Gauss-Hermite quadrature nodes and weights to compute the above integrals for each observation. Notice also that the nodes and weights are independent of the observation because we always integrate with respect to the same distribution (standard normal). This is a good news because we can compute the nodes and weights only once for all the iterations. Moreover, the Gauss-Hermite quadrature requires very few nodes for very high precision, usually 21 is enough. Therefore, if we have n observations of the first type, and we are using m nodes, we need to evaluate the integrand $n\times m$ times in order to compute $P\left(2E|x,y\right)$ for all the observations. This should be quite fast because the function Φ is fast and we have analytical solution to the model. The resulting object $P\left(2E|x,y\right)$ is $n\times 1$ vector.

Now we show how to treat the observations with only one of the wages observed. Suppose that we have n observations of the form

Obs. #	x	y
1	x^1	0
2	x^2	0
3	x^3	0
$\vdots$	$\vdots$	$\vdots$

We are interested in computing $P\left(1M|x\right)$, which is a $n\times 1$ vector. Now we need to integrate twice; once with respect to z and once with respect to y, as shown below

$$\begin{aligned}P\left(1M|x,y\right)&=\int\Phi\left(z+\frac{V_{1M}\left(x,y\right)-V_{2E}\left(x,y\right)}{\sigma}\right)\cdot\Phi\left(z+\frac{V_{1M}\left(x,y\right)-V_{1F}\left(x,y\right)}{\sigma}\right)\phi\left(z\right)dz\\P\left(1M|x\right)&=\int P\left(1M|x,y\right)f\left(y|x\right)dy\end{aligned}$$

The problem is that the objective function is not smooth with respect to y, so we might need many nodes for the integration with respect to y. In other words, we need to compute

$$\int\left[\int\Phi\left(z+\frac{V_{1M}\left(x,y\right)-V_{2E}\left(x,y\right)}{\sigma}\right)\cdot\Phi\left(z+\frac{V_{1M}\left(x,y\right)-V_{1F}\left(x,y\right)}{\sigma}\right)f\left(y|x\right)dy\right]\phi\left(z\right)dz$$

Notice that

$$P(1M|x,y,z) = \Phi\left(z + \frac{V_{1M}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(z + \frac{V_{1M}(x,y) - V_{1F}(x,y)}{\sigma}\right)$$

Thus, we need to compute

$$P(1M|x) = \int \left[\int P(1M|x,y,z) f(y|x) dy\right] \phi(z) dz$$

and

$$\int P(1M|x,y,z) f(y|x) dy = \int P\left(1M|x, s\cdot\sigma_{Y|X=x} + \mu_{Y|X=x}, z\right) \phi(s) ds$$

Finally, the treatment of the observations of the third kind is the same as the second. We have observations of the form

Obs. #	x	y
1	0	y^1
2	0	y^2
3	0	y^3
$\vdots$	$\vdots$	$\vdots$

and we need to compute

$$\begin{aligned} P(1F|x,y) &= \int \Phi\left(z + \frac{V_{1F}(x,y) - V_{2E}(x,y)}{\sigma}\right) \cdot \Phi\left(z + \frac{V_{1F}(x,y) - V_{1M}(x,y)}{\sigma}\right) \phi(z) dz \\ P(1F|y) &= \int P(1F|x,y) f(x|y) dx \end{aligned}$$

Here again we have a double integral.

In some sense, the Maximum Likelihood approach is simpler than the Method of Moments since we don't have that many integrals in the objective function. Moreover, and this is the main advantage, we are not required to compute numerically the limits between the $2E$, $1M$, and $1F$ regions. In this model there are no such regions. Any couple can be anywhere because of the $\varepsilon' s$.

7 CES utility

In this section we deviate from the log utility and and try the CES utility. There are two reasons why in our model changes in taxes do not generate big changes in participation: (1) the taxes did not change enough, and (2) our choice of utility function does not allow big responces of participation to changes in taxes.

The household preferences are

$$u(c,l) = \left[\alpha c^{\rho} + (1-\alpha) l^{\rho}\right]^{1/\rho}, \quad \rho \leq 1$$

The 2-earner household problem is

$$
\begin{aligned}
& V_{2E}\left(w_m, w_f\right) = \\
& \max_{c_m, c_f} \left\{ \lambda \left[\alpha c_m^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho} + (1-\lambda)\left[\alpha c_f^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho} \right\} \\
& \text{s.t.} \quad c_m + c_f = \frac{w_m + w_f - T\left(w_m, w_f\right)}{1+\tau} \equiv I_{2E}
\end{aligned}
$$

F.O.C.

$$
\begin{aligned}
\lambda \frac{1}{\rho}\left[\alpha c_m^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho-1} \alpha\rho c_m^{\rho-1} &= (1-\lambda)\frac{1}{\rho}\left[\alpha c_f^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho-1} \alpha\rho c_f^{\rho-1} \\
\lambda \left[\alpha c_m^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho-1} c_m^{\rho-1} &= (1-\lambda)\left[\alpha c_f^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho-1} c_f^{\rho-1} \\
\lambda \left[\alpha c_m^{\rho} c_m^{-\rho} + (1-\alpha)(1-h_m)^{\rho} c_m^{-\rho}\right]^{\frac{1-\rho}{\rho}} &= (1-\lambda)\left[\alpha c_f^{\rho} c_f^{-\rho} + (1-\alpha)(1-h_f)^{\rho} c_f^{-\rho}\right]^{\frac{1-\rho}{\rho}} \\
\lambda^{\frac{\rho}{1-\rho}}\left[\alpha + (1-\alpha)(1-h_m)^{\rho} c_m^{-\rho}\right] &= (1-\lambda)^{\frac{\rho}{1-\rho}}\left[\alpha + (1-\alpha)(1-h_f)^{\rho} c_f^{-\rho}\right] \\
\lambda^{\frac{\rho}{1-\rho}}\alpha + \lambda^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_m)^{\rho} c_m^{-\rho} &= (1-\lambda)^{\frac{\rho}{1-\rho}}\alpha + (1-\lambda)^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_f)^{\rho} c_f^{-\rho} \\
\lambda^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_m)^{\rho} c_m^{-\rho} &= \alpha\left((1-\lambda)^{\frac{\rho}{1-\rho}} - \lambda^{\frac{\rho}{1-\rho}}\right) + (1-\lambda)^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_f)^{\rho} c_f^{-\rho} \\
c_m^{-\rho} &= \frac{\alpha\left((1-\lambda)^{\frac{\rho}{1-\rho}} - \lambda^{\frac{\rho}{1-\rho}}\right)}{\lambda^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_m)^{\rho}} + \left(\frac{(1-\lambda)^{\frac{\rho}{1-\rho}}(1-h_f)^{\rho}}{\lambda^{\frac{\rho}{1-\rho}}(1-h_m)^{\rho}}\right) c_f^{-\rho}
\end{aligned}
$$

This can be written as

$$
c_m^{-\rho} = a + b c_f^{-\rho}
$$

where

$$
\begin{aligned}
a &= \frac{\alpha\left((1-\lambda)^{\frac{\rho}{1-\rho}} - \lambda^{\frac{\rho}{1-\rho}}\right)}{\lambda^{\frac{\rho}{1-\rho}}(1-\alpha)(1-h_m)^{\rho}} \\
b &= \frac{(1-\lambda)^{\frac{\rho}{1-\rho}}(1-h_f)^{\rho}}{\lambda^{\frac{\rho}{1-\rho}}(1-h_m)^{\rho}}
\end{aligned}
$$

Thus, the solution to the household's problem is obtained by solving the following system:

$$
\begin{aligned}
c_m^{-\rho} &= a + b c_f^{-\rho} \\
c_m + c_f &= I_{2E}
\end{aligned}
$$

Substituting the budget constraint gives

$$
c_m^{-\rho} = a + b\left(I_{2E} - c_m\right)^{-\rho}
$$

This equation has analytical solution if $a = 0$, or $\lambda = 0.5$. In this case we have

$$\begin{aligned}
c_m^{-\rho} &= b\left(I_{2E} - c_m\right)^{-\rho} \\
c_m &= b^{-1/\rho}\left(I_{2E} - c_m\right) \\
c_m &= \left(\frac{(1-h_f)^{\rho}}{(1-h_m)^{\rho}}\right)^{-1/\rho}\left(I_{2E} - c_m\right) \\
c_m &= \left(\frac{(1-h_m)^{\rho}}{(1-h_f)^{\rho}}\right)^{1/\rho}\left(I_{2E} - c_m\right) \\
c_m &= \left(\frac{1-h_m}{1-h_f}\right)\left(I_{2E} - c_m\right) \\
c_m + \left(\frac{1-h_m}{1-h_f}\right)c_m &= \left(\frac{1-h_m}{1-h_f}\right)I_{2E} \\
c_m\left[1 + \frac{1-h_m}{1-h_f}\right] &= \left(\frac{1-h_m}{1-h_f}\right)I_{2E} \\
c_m\left[\frac{1-h_f+1-h_m}{1-h_f}\right] &= \left(\frac{1-h_m}{1-h_f}\right)I_{2E} \\
c_m &= \frac{(1-h_m)\,I_{2E}}{(1-h_m+1-h_f)} \\
c_f &= \frac{(1-h_f)\,I_{2E}}{(1-h_m+1-h_f)}
\end{aligned}$$

Notice that $1 - h_m$ is the male's leisure and $1 - h_f + 1 - h_m$. The above condition says that male's consumption as a fraction of total consumption is the same is his leisure out of the total leisure. We introduce a shorthand notation for the above demands

$$\begin{aligned}
c_m &= \phi I_{2E} \\
c_f &= (1-\phi)\,I_{2E}
\end{aligned}$$

Substituting in the utility gives

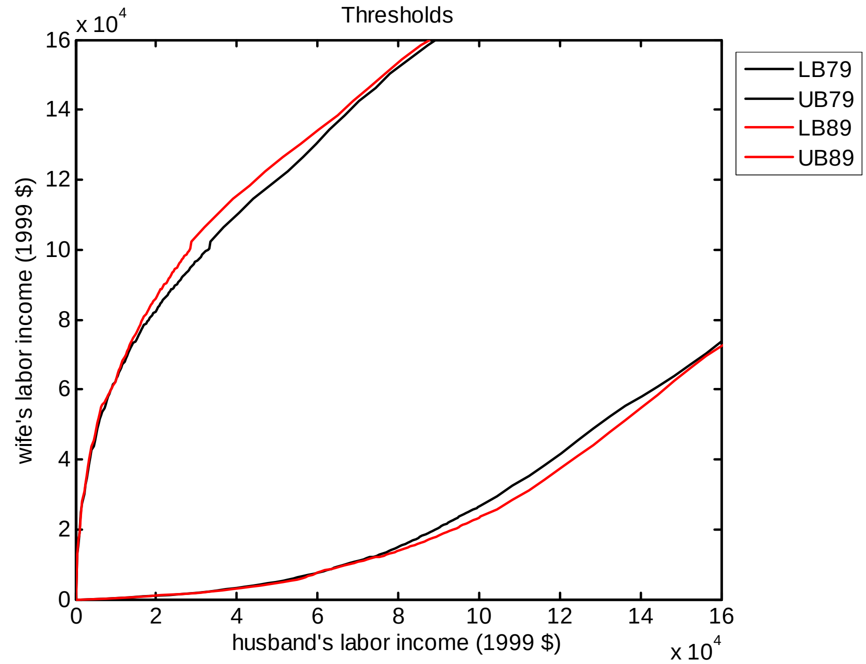
$$\begin{aligned}
V_{2E} &= \left[\alpha\left(\phi I_{2E}\right)^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho} + \left[\alpha\left((1-\phi)\,I_{2E}\right)^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho} \\
V_{1M} &= \left[\alpha\left(\phi I_{1M}\right)^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho} + \left[\alpha\left((1-\phi)\,I_{1M}\right)^{\rho} + 1-\alpha\right]^{1/\rho} \\
V_{1F} &= \left[\alpha\left(\phi I_{1F}\right)^{\rho} + 1-\alpha\right]^{1/\rho} + \left[\alpha\left((1-\phi)\,I_{1F}\right)^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho}
\end{aligned}$$

The household will choose to be a 2-earner couple instead of a 1-earner male couple if

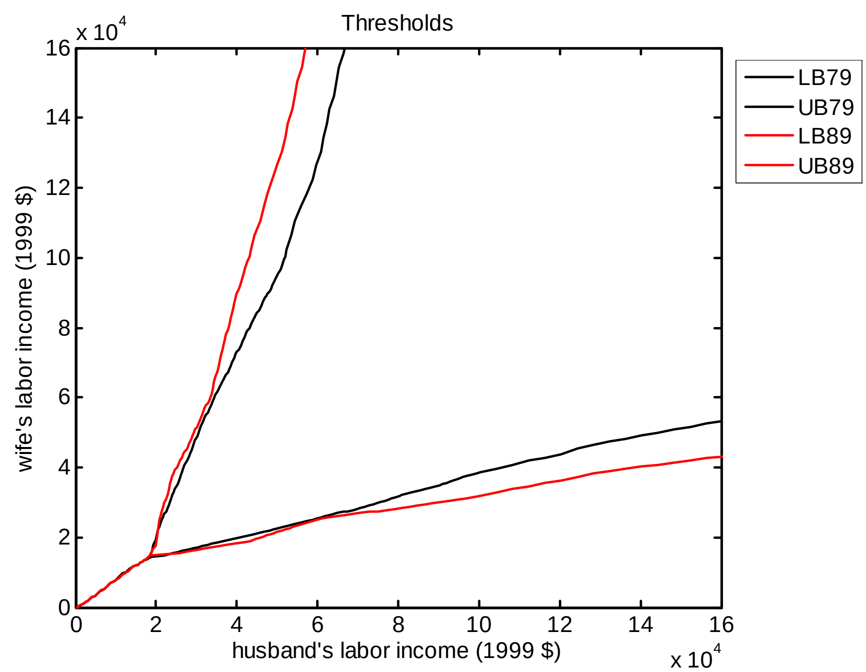
$$\begin{aligned}
& \left[\alpha\left(\phi I_{2E}\right)^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho} + \left[\alpha\left((1-\phi)\,I_{2E}\right)^{\rho} + (1-\alpha)(1-h_f)^{\rho}\right]^{1/\rho} \\
\geq\; & \left[\alpha\left(\phi I_{1M}\right)^{\rho} + (1-\alpha)(1-h_m)^{\rho}\right]^{1/\rho} + \left[\alpha\left((1-\phi)\,I_{1M}\right)^{\rho} + 1-\alpha\right]^{1/\rho}
\end{aligned}$$

7.1 The effect of ρ on the thresholds

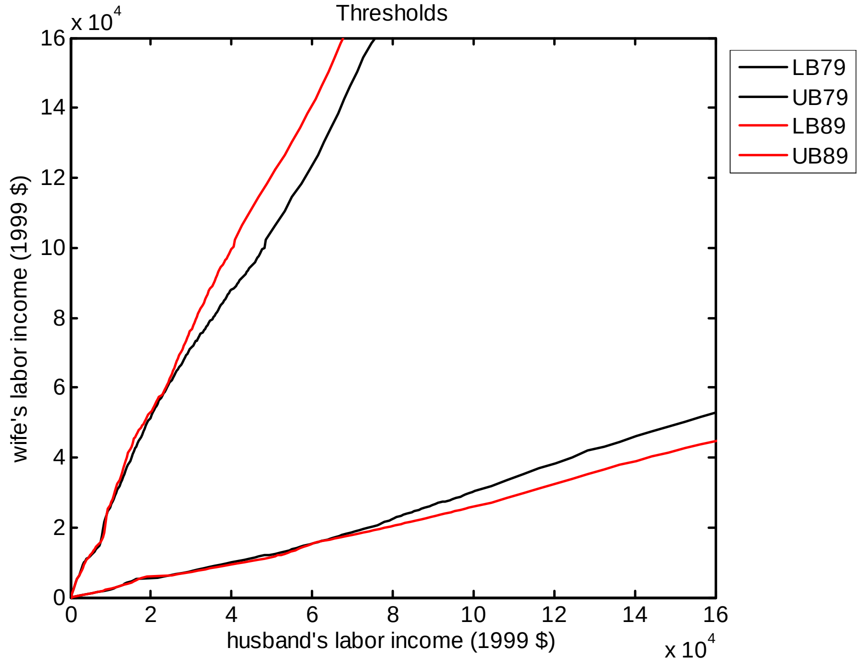
The next graph shows the thresholds for $\rho = -1$ (elasticity of substitution is $1/(1-\rho) = 0.5$), and $\alpha = 0.999985$.



The next graph shows the thresholds for $\rho = 0.5$ (elasticity of substitution is $1/(1-\rho) = 2$), and $\alpha = 0.004$.



The next graph shows the thresholds for $\rho = 0$ (elasticity of substitution is $1/(1-\rho) = 1$), and $\alpha = 0.5$.



Based on the above graphs, it doesn't seem that the week response of the the model to changes in taxes is rooted in the choice of utility. Observe that the elasticity of substitution between consumption and leisure affects the shape of the thresholds, but the response of the thresholds to changes in taxes remains small. Moreover, it seems that when the elasticity of substitution is low, the participation of wives married to low income husbands is too high, and when the elasticity of substitution is high, the fraction of 2-earner couples becomes too low for couples with low income of the husband.

7.2 Labor supply elasticity

One way to measure the labor supply elasticity in our model is as follows. First, define the labor supply of male and female by

$$\begin{aligned} L_m &= P(1M \cup 2E) \cdot h_m \\ L_f &= P(1F \cup 2E) \cdot h_f \end{aligned}$$

Let the wages be LogNormally distributed

$$(w_m, w_f) \sim LN(\mathbf{m}, \mathbf{S})$$

Then, the aggregate labor supply elasticities of male and female are

$$\begin{aligned} \eta_m &\equiv \frac{\%\Delta L_m}{\%\Delta m_1} = \frac{\%\Delta P(1M \cup 2E)}{\% m_1} = \frac{\partial P(1M \cup 2E)}{\partial m_1} \frac{m_1}{P(1M \cup 2E)} \\ \eta_f &\equiv \frac{\%\Delta L_f}{\%\Delta m_2} = \frac{\%\Delta P(1F \cup 2E)}{\%\Delta m_2} = \frac{\partial P(1F \cup 2E)}{\partial m_2} \frac{m_2}{P(1F \cup 2E)} \end{aligned}$$

That is, we simply measure the percentage change in the fraction of male and female that work, as we change the means of their respective wages by 1%. We could also consider labor supply elasticities for specific types of couples, for example, by intervals of husband's income. Then we have

$$\begin{aligned}\eta_m(a,b) &= \frac{\%\Delta P\left(1M \cup 2E | a \leq w_m \leq b\right)}{\%\Delta m_1} \\ \eta_f(a,b) &= \frac{\%\Delta P\left(1F \cup 2E | a \leq w_m \leq b\right)}{\%\Delta m_2}\end{aligned}$$

In the accounting paper, the labor supply elasticity of male is 0, since all the male are working full time regardless of the wages. We can however report the labor supply elasticity of female as

$$\eta_f(a,b) = \frac{\%\Delta P\left(2E | a \leq w_m \leq b\right)}{\%\Delta m_2}$$

7.3 Labor supply elasticity in the static labor supply model

This section contains the analysis of the static model with consumption and leisure choice. The household preferences are:

$$u(c,l) = \left[\alpha c^{\rho} + (1-\alpha) l^{\rho}\right]^{1/\rho}, \quad \rho \leq 1$$

The household's problem is

$$\begin{aligned} & \max_{c,h} \left[\alpha c^{\rho} + (1-\alpha)(1-h)^{\rho}\right]^{1/\rho} \\ & s.t. \\ c &= hw \end{aligned}$$

F.O.C.

$$\begin{aligned}\frac{\rho(1-\alpha)(1-h)^{\rho-1}}{\rho\alpha c^{\rho-1}} &= w \\ \frac{(1-\alpha)(1-h)^{\rho-1}}{\alpha} &= wc^{\rho-1}\end{aligned}$$

Substituting the budget constraint

$$\begin{aligned}\frac{(1-\alpha)(1-h)^{\rho-1}}{\alpha} &= w(hw)^{\rho-1} \\ \left(\frac{1-\alpha}{\alpha}\right)\left(\frac{1-h}{h}\right)^{\rho-1} &= w^{\rho} \\ \left(\frac{1-\alpha}{\alpha}\right)^{\frac{1}{\rho-1}}\left(\frac{1-h}{h}\right) &= w^{\frac{\rho}{\rho-1}}\end{aligned}$$

Recall that the elasticity of substitution between consumption and leisure is

$$\sigma = \frac{1}{1-\rho}$$

Thus

$$1-\sigma = 1-\frac{1}{1-\rho} = \frac{1-\rho-1}{1-\rho} = \frac{\rho}{\rho-1}$$

The demand then can be written as

$$\begin{aligned}
\left(\frac{1-\alpha}{\alpha}\right)^{-\sigma}\left(\frac{1-h}{h}\right) &= w^{1-\sigma} \\
1-h &= h\left(\frac{1-\alpha}{\alpha}\right)^{\sigma} w^{1-\sigma} \\
1-h &= h\left(\frac{1-\alpha}{\alpha}\right)^{\sigma} w^{1-\sigma} \\
h &= \frac{1}{1+\phi w^{1-\sigma}}
\end{aligned}$$

where

$$\phi = \left(\frac{1-\alpha}{\alpha}\right)^{\sigma}$$

The labor supply elasticity in this model is

$$\eta = \frac{dh}{dw}\frac{w}{h} = -\frac{\phi\,(1-\sigma)\,w^{-\sigma}}{(1+\phi w^{1-\sigma})^2}\frac{w}{h} = -\frac{\phi\,(1-\sigma)\,w^{1-\sigma}}{h\,(1+\phi w^{1-\sigma})^2}$$

Notice that in the log utility case ($\rho = 0$, $\sigma = 1$), the labor supply is perfectly inelastic. That is, we get the familiar result that in the Cobb-Douglas case the labor supply is independent of the wage. If consumption and leisure are substitutes (i.e., $\rho > 0$, $\sigma > 1$), the labor supply ilasticity is positive. This means that when w goes up, the labor supplied will go up. Finally, if consumption and leisure are complements (i.e., $\rho < 0$, $\sigma < 1$), then the labor supply elasticity is negative. This means that when w goes up, the labor supplied goes down. To summarize, in this model there is a direct link between the elasticity of substitution σ and the labor supply elasticity.

Our model is different from the one just described in that it has descrete choice of participation and the utility is that of a couple instead of an individual. Therefore, in our model there is no direct link between the elastcity of substitution and the elasticity of labor supply. Changing ρ changes the shape of the thresholds in our model, but it is not clear how that affects the magnitude of impact of changes in wages on the labor supply.

References

[1] Green, William H., *Econometric Analysis, 4th adition,* Prentice Hall, 2000.

[2] Averill M. Law and W. David Kelton, *Simulation Modeling and Analysis, 3rd edition, McGraw-Hill Higher Education, 2000.*

[3] Miranda, J. Mario, Paul L. Fackler (2002), "Applied Computational Economics and Finance", *MIT press.*