1. Introduction

1.1. What is the study of Labor Economics about?

An examination of how the labor market works,

-What affects the unemployment rate, and how?

-What affects the labor participation rates, and how?

An examination of how employers & employees interact.

-What are the consequences on human capital accumulation?

-What are the consequences of marital patterns in the labor market?

An examination of how agents react to, and shape social rules and forces, and hence affecting their choices.

1.2. What are some pertinent questions we can hope to answer along the way?

1. Why did labor force participation of women rose over the past century?
2. Do minimum wages increase unemployment rate of less-skilled workers?
3. What is the impact of workplace regulation on employment and earnings?
4. Are government subsidies of investments in human capital accumulation an effective means of raising worker welfare, and which workers?
5. What is the impact of affirmative action programs on the earnings of women and minorities?
6. What is the impact of unions on members, their offspring, and other members of the economy?
7. Do generous unemployment insurance benefits lengthen the duration of spells of unemployment?
8. Who really benefits from subsidies in post secondary or tertiary education?

9. What are the consequences of "liberal" divorce laws on the marriage markets, such as labor force participation, and investment in children, and the type of marriages that take place?

10. Did improved contraception technology raise the incidence of common law partnerships, and hence marriage rates?

1.3. Who are the actors in the Labor Market

1.3.1. Labor chooses;

1. How much education to attain?

2. Whether to work or not?

3. To work for whom and for how much?

4. How much effort to allocate into a work?

5. Which skills to acquire on the job?

6. When to quit and seek new employement?

7. Whether to join a labor union?

8. Who to choose for a spouse?

9. How much to invest in own children's welfare and education?

What does all these mean? $\Rightarrow$ maximize utility function or own welfare$\rightarrow$ choice for consumption, spouse, investment in children, hours of work per day/week.

The summation of all these decisions or choices gives demand for goods, marriage market equilibrium, equilibrium outcome of children, and labor supply function/curve (See figure 1).

1.3.2. Firms chooses;

What type of workers to hire?

How many workers to hire?

Duration of work week?

How much capital to augment the use of labor?

How much training to offer the workers?

How to ensure workers maximize their effort and do not shirk their respnsibility?

The firm's choice for labor is dependent on the demand of the good they produce, so labor demand is a derived demand. The variables are chosen to maximize the firms profit. The summation of the labor decision for all firms gives the labor demand function/curve (See figure 1).

1.3.3. Equilibrium & the Government;

As workers search for jobs, and firms search for workers, the conflicting interests balances out, and generates the labor market equilibrium (See figure 1).

However, if the types of workers generated by the market does not meet the needs of the economy or society, or if there is in consequence, and unacceptable level of unemployment from the perspective of aggregate welfare of the society, there is a role for the Government.

Some tools available to the government are,

1. Impose taxes on workers' earning to meet other public needs.
2. Subsidize training in profession that the economy or society has a shortage for either currently or future.
3. Impose hiring rules based on ethnic makeup of labor force in the firm or economy.
4. Alter immigration rules to augment immediate labor market needs.

5. Legislate laws to achieve desire marriage market welfare or of specific populace, example children from divorced families.

It thus provides rules under which all the actors in the labor market work through. Example?

1.4. Difference between Positive and Normative Economics

2. Appendix:

2.1. Regression Analysis:

Much of Labor Economics is about the estimation of economic relationships, and the examination of whether theoretical predictions match those estimated relationships. It thus makes extensive use of econometrics. We are concerned with both the sign and size of impact. Although we will not be using much econometric analysis, a basic understanding of what it means to perform a regression analysis, and the ensuing concerns would help you understand the benefits and limits of the results, and allow you to understand contemporary work in Labor Economics. The more advance or inquisitive student may want to consult

1. J.M. Wooldridge (1999) "Introductory Econometrics: A Modern Approach", South-Western Publishing

2. J. Johnston and J DiNardo (1996) "Econometric Methods", McGraw-Hill/Irwin

Example 1 *Let's focus on a single question: Why do some occupations pay more than others?*

Figure 2 provides a possible manifestation of observable relationship between wages and education from data sources such as the Canadian Labour Market Activity Survey (LMAS). What is one obvious factor that affects our earnings? There is obviously several, but let's *focus on educational attainment. Then a possible way to characterize this relationship in an estimatible equation is*

$$w = \alpha + \beta s \tag{1}$$

where w=wages (the dependent variable), and s=educational attainment (independent variable). The aim is to obtain numrical estimates of α and β from say the LMAS, or to fit a line that best approximates the true relationship (The bold line in figure 2). But there are obviously idiosyncratic differences between individuals which explains why dots which identifies each individual does not correspond to the line, i.e be on the line. A typical estimation equation is then

$$w_i = \alpha + \beta s_i + \epsilon_i \tag{2}$$

where ϵ_i is the portion of the relationship that is idiosyncratic, or due to some systematic relationship with the other variables that may explain pay differential across occupations. For the moment, let us assume that education fully explains differences in wages. Then by regression analysis, we are in effect trying to fit the equation (2) that minimize these idiosyncratic differences and prevent it from distracting us from the true relationship. Let there be N observations, then we have to solve

$$\begin{aligned} &\min_{\alpha,\beta} \sum_{i=1}^{N} (\epsilon_i)^2 \\ \Rightarrow \quad &\min_{\alpha,\beta} \sum_{i=1}^{N} (w_i - \alpha - \beta s_i)^2 \end{aligned}$$

Why do we $\min \sum (\epsilon)^2$ *and not* $\min \sum \epsilon$[1] *?*

[1] The solution to the above problem for the estimate of α, and β is,

$$\begin{aligned} 2\left(\sum_{i=1}^{N} w_i - N\widehat{\alpha} - \widehat{\beta}\sum_{i=1}^{N} s_i\right) &= 0 \\ 2\left(\sum_{i=1}^{N} w_i s_i - \widehat{\alpha}\sum_{i=1}^{N} s_i - \widehat{\beta}\sum_{i=1}^{N} (s_i)^2\right) &= 0 \end{aligned}$$

Regression requires that ϵ_i be distributed under a normal distribution with,

$$\begin{aligned} E(\epsilon_i) &= 0 \\ Var(\epsilon_i) &= \sigma^2 \\ Cov(\epsilon_i, \epsilon_j) &= 0 \end{aligned}$$

or simply put $\epsilon_i \tilde{} N(0, \sigma^2)$. There are other assumptions the interested reader may easily discover.

α is just the intercept of the line in figure 2, and β is just the slope of this line. So what? What does the intercept and slope of the relationship mean? The intercept gives gives the lowest wage that could be had if an agent has not had any education or schooling. Since the slope in upward sloping, it means that the greater the number of years of schooling (increase along the horizontal axis), the greater would be the wage possible (the increase along the vertical axis). Or differentiating w in (1) with respect to s,

$$\frac{dw}{ds} = \frac{\textit{1 unit change in wages}}{\textit{1 unit change number of years of schooling}} = \beta$$

which says exactly the same thing[2]*. But is this a good representation of the relationship. Is that a better way to represent this relationship. Consider the following where instead of*

Solving this two equation obtained from the first order condition for α and β respectively, we obtain the respective solution for the estimates as

$$\begin{aligned} \widehat{\alpha} &= \frac{\sum_{i=1}^{N} w_i}{N} - \widehat{\beta}\frac{\sum_{i=1}^{N} s_i}{N} \\ &= \overline{w} - \widehat{\beta}\overline{s} \end{aligned}$$

Substituting this into (??)

$$\begin{aligned} \widehat{\beta} &= \frac{\sum_{i=1}^{N} w_i s_i - \overline{w}\sum_{i=1}^{N} s_i}{\sum_{i=1}^{N} (s_i)^2 - \overline{s}\sum_{i=1}^{N} s_i} \\ &= \frac{N\sum_{i=1}^{N} w_i s_i - \sum_{i=1}^{N} w_i \sum_{i=1}^{N} s_i}{N\sum_{i=1}^{N} (s_i)^2 - \left(\sum_{i=1}^{N} s_i\right)^2} \end{aligned}$$

[2]Mathematics is just a shorthand, not a new language.

having an equation that relates wages and year of education,

$$\log w = \alpha + \beta s \tag{3}$$

That is we substitute the logarithm of wages instead of wages. Why is this a better representation? Although wages in itself is a good measure, how great is the jump in wages for each additional year of schooling? Differentiate w in (3) with respect to s,

$$\frac{d\log w}{ds} = \beta = \frac{1}{w}\frac{dw}{ds} = \frac{\frac{dw}{w}}{ds} = \frac{\textit{Percentage change in wages}}{\textit{1 unit change in the number of years of schooling}}$$

The estimated equation is now

$$\widehat{w}_i = \widehat{\alpha} + \widehat{\beta} s_i$$

Then the error for each observation is

$$w_i - \widehat{w}_i = \widehat{\epsilon}_i$$

Then the estimated variance, $\widehat{\sigma}^2$ *is*

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{N} \widehat{\epsilon}_i^2}{N-2}$$

The denominator is $N-2$ *because we lose 2 degrees of freedom from the estimation of* α *and* β*. Using this we may calculate the variance of the estimates of* α*, and* β*.*

$$Var\left(\widehat{\beta}\right) = \frac{\sigma^2}{\sum_{i=1}^{N}\left(s_i - \overline{s}\right)^2} \tag{4}$$

$$Var\left(\widehat{\alpha}\right) = \frac{\sum_{i=1}^{N} s_i^2}{N\sum_{i=1}^{N}\left(s_i - \overline{s}\right)^2}\sigma^2 \tag{5}$$

Where we use $\widehat{\sigma}^2$ *if* σ^2 *is unknown. Both equations (4) and (5) gives us an idea how good our estimates are, in the sense of whether the variances are large. If its large, it means there is too much variation in the data that is not explained by the explanatory variable. But how* *do we know if a variance is too large, or just right? The answer is in the assumption that* ϵ_i *is normally distributed. Consequently, so is the distribution of our estimates, that is*

$$\widehat{\beta}\tilde{\ }N\left(\beta, Var\left(\widehat{\beta}\right)\right)$$
$$\widehat{\alpha}\tilde{\ }N\left(\alpha, Var\left(\widehat{\alpha}\right)\right)$$

Next note that,

$$\frac{\widehat{\beta}-\beta}{\sqrt{Var\left(\widehat{\beta}\right)}}\tilde{\ }N\left(0,1\right)$$
$$\frac{\widehat{\alpha}-\alpha}{\sqrt{Var\left(\widehat{\alpha}\right)}}\tilde{\ }N\left(0,1\right)$$

That is they have a standardized normal distribution, with a mean of 0 and a variance of 1 (See figure 3) if σ^2 *is known, and the statistic is known as the Z statistic. If* σ^2 *is unknown, we replace* σ^2 *with* $\widehat{\sigma}^2$*, and the above becomes know as a student t distribution, the shape of the distribution is similar to a normal distribution of figure 3, and the statistic is called the t-statistic. But what is the hypothesis we wish to test, or what is reported typically in journal articles? We typically want to know if there is infact a relationship (Only after there is some strong evidence there is something we wish to explain, do we seek to explain it with a theoretical exposition). If we are incorrect, that is there is no relationship, then* $\beta = 0$*. We say this is the null hypothesis we wish to test, and write* $H_0 : \beta = 0$*. A priori, we may not know if* β *is either positive or negative, and so we typically test* H_0 *against the alternative hypothesis that* β *is infact none zero,* $H_1 : \beta \neq 0$*. If our* $\widehat{\beta}$ *is indeed zero, then the t statistic should not be very far from the truth, i.e. 0. But if it is non-zero, then the t statistic must be quite large, which occurs when* $\widehat{\beta}$ *is large, and its variance is relatively small. To minimize the probability of us making the incorrect conclusion that* β *is not zero when it actually is, we set the level we could accept, that is the probability that the statistic is at the ends of the distribution (As in figure 3), at typically 5%. The cutoff statistic, as a rule of thumb is*

2. Then if the t-statistic is greater or equal to 2, we say it is statistically significant, and otherwise statistically insignificant. Another quick way is to see if the coefficient estimated is at least twice as large as its standard errors.

Even if our estimates are significant, it is possible that errors or disperson is substantial. That is there is still susbtantial variation in the data that our estimation equation cannot explain. This is the norm. So in our example, some of the variation or dispersion could be due to gender differences (due to gender discrimination, differential fecundity), parent's social status as measured by their income and education, ethnicity, immigration status,... All or some of which may explain any amount of dispersion we see between the actual observations and our regression line.

What is the generalization of the basic regression?

$$\log w = \alpha + \beta_1 s + \beta_2 g + \beta_3 ps + \beta_4 r$$

where g is for gender (where g=1 if say female, and 0 otherwise), ps is for parent's income, and r is for race (where r=1 if caucasian, 0 otherwise). We will broach more issues as they reveal themselves. (You should also consider other limitations of regression analysis. What if the independent variables are correlated with each other?)

3. Class Project:

1. Empirical studies in the US has found that increasing direct costs of tertiary education may reduce enrollments, and it disadvantages individuals from poorer social circumstance. There is however dispute as to whether this greater response among the disadvantaged is evidence that credit constraints are more binding (Kane, 1994). Carneiro and Heckman (2002), have infact argued that there is little evidence of important credit constraints affecting post-secondary enrollments, consequently allaying any concerns for the adverse consequences from fee increases. Do you agree that alter-

ations to direct costs of post secondary education has no impact on the disadvantaged? Examine arguements and evidence from all parties, highlighting their concerns (both theoretical and empirical), as well as your own. Explain how you might contribute or resolve the debate, and what type of data set would you require.

2. There is extensive empirical research on the positive impact that unionization has on different aspects of industrial relations, much of it focused on wage differences (Card, 1996; Fang and Verma, 2002), productivity (Gregg, Machin and Metcalf, 1993), and reducing turnover (Elias, 1994). However, there has been no work performed on the income and educational mobility among families with unionized members. To be precise, do you think by virtue an individual is a member of a union affects how much, and how he invest in human capital accumulation for his children, and consequently the income of his children as they enter the labor force. If so, why and what is the probable mechanism? If not, why? If you had data (what kind of data would you need), how would you prove or disprove your hypothesis?

3. Canada introduced the point system in 1967 for its immigrant selection policy in an attempt to eliminate the bias towards accepting migrants of European decent. It has been argued that this selection policy allows Canada to select "winners" who are better able to assimilate into the economy, and society, and contribute economically to the vibrancy of Canada. How would verify these statements, and with what kind of data? Do you think these policies will have an effect on the marriage market in Canada? If not, why not? What effects do you think these policies have on the skills or educational levels in Canada, compared to U.S? How would the labor force evolve, if there are no marriage market effects, or when there are?

4. It has been suggested that with the change in composition of migrants into Canada, the assimilation rates has slowed substantially. Some has argued that this newer cohorts do not possess the same linguistic abilities, and further as a result of cultural differences are

less able to assimilate into contemporary Canadian society. What are some possible problems that may prevent you from identifying the reason for this phenomonon? How would you rectify these deficiency? What would the perfect dataset have to be? Why do you think the assimilation rates are lower, disregarding the possibility of discrimination?

5. A common characterization of the economic outlook of the maritime and the prairie provinces in Canada are that they are laguishing in poverty, with substantially slower growth rates that the provinces. Do you think the slower growth rates is a result of a brain drain to the larger urban centres such as Toronto, Montreal or Vancouver where there are greater opportunities? Or do you think the brain drain or outward migration is a result of the faster growth rates in the other provinces? This is the classic question about causality and endogeneity. How would you solve this problem in estimation? What do you think are the consequences of this outward migration on the quality of labor force in the maritime and prairie provinces? What do you think are the effects on the marriage market, that is the quality of matches that take place in these poorer provinces, the quality of children in terms of educational attainment and abilities (if you believe abilities are largely heriditary)? What does this mean in terms of future growth potential?

6. Question of your choice which must be approved by myself?