Final Exam ECON3715/4715 – Labour Economics
Autumn 2018

This exam has 5 questions, with 16 sub-questions. Each sub-question counts equally. When answering the questions on the exam you should be brief and to the point! Make sure to write clearly. Difficult to decipher answers will not be counted!

---

1. In this question you have to indicate whether you think the statement is true or false and explain why. You do not get any points if you only state whether the statement is true or false.

   (a) The intertemporal substitution of labor supply measures the change in hours worked from a permanent increase in wages.

   False. The intertemporal substitution of labor supply measures the change in hours worked from an *anticipated* increase in wages along the life-cycle wage profile of an individual (also known as evolutionary wage change). Since an anticipated wage increase doesn't change the life-cycle wage profile of an individual known to the agent at start of the career, the marginal utility of lifetime wealth is kept constant. In contrast, a permanent (unanticipated) increase in wages will shift the life-cycle wage path and induce wealth effects.

   (b) A firm has more market power when slope of the labor supply curve facing a firm is higher.

   True. If the slope of the labor supply curve facing a firm is high, the elasticity of labor supply facing a firm is low and the firm (a non-discriminatory monopsonist) can set a wage below the competitive wage rate without loosing many workers. The higher the slope of the firm-level labor supply curve, the larger is the wedge between the value of marginal product of labor (VMP) and the monoposony wage rate, and the more market power does the firm hold. In comparison, with a competitive labor market, the elasticity of labor supply facing a single firm is infinite and a firm can't set a wage below the competitive wage rate without loosing all workers.

(c) Use the information below showing the present value of lifetime productivity and the costs of schooling for high- and low-productivity workers to answer.

Type of worker	**Present value of lifetime productivity**	**Cost of a year of schooling**
low-productivity	450 000	20 000
high-productivity	600 000	10 000

If high-productivity workers obtain 7 years of schooling, a separating equilibrium occurs where every worker is paid his or her present value of lifetime productivity.

False. In a separating equilibrium in which workers are paid their present value of lifetime productivity, it must be the case that high productivity workers obtain a number of years of schooling for which it is “unprofitable” for low-productivity workers to obtain the same number of years of schooling.

$$600000 - 10000 \cdot y \geq 450000 \longrightarrow y \leq 15$$

$$600000 - 20000 \cdot y < 450000 \longrightarrow y > 7.5$$

In a separating equilibrium high-productivity workers obtain $7.5 < y \leq 15$ years of schooling. If high productivity obtain 7 years of schooling, it would be profitable for low productivity workers to mimick high-productivity workers and this would thus result in a pooling equilibrium.

(d) Consider two individuals with different discount rates who are otherwise identical. Both individuals are unemployed and searching for jobs. We expect the individual with a higher discount rate to have a longer unemployment duration.

False. If an individual has a higher discount rate he cares more about the present than the future and this decreases the present value of the expected wage increase. It decreases the asking wage

$$\widetilde{w} = b - c + P(w \geq \widetilde{w}) \times \frac{E(w - \widetilde{w}|w \geq \widetilde{w})}{r}$$

and the individual will accept job offers with lower wage offers and the expected unemployment duration will therefore be shorter.

2. Suppose that individuals select between working in sector A and working in sector B. An individual can only work in one of these two sectors. Individual i can produce X_i^s goods with unit price P^s and thus earn income $W_i^s = X_i^s P^s$ in sector $s = \{A, B\}$. Imagine that sectoral productivities are perfectly correlated for all individuals such that the following relationship holds:

$$\log X_i^A = \phi \log X_i^B.$$

Suppose that individuals are income-maximizing, i.e., sector B is selected if $\log W_i^B > \log W_i^A$ and sector A is selected if otherwise.

(a) Suppose $\phi \in (0, 1)$. Do workers with the highest productivity in sector B decide to work in sector B? And do workers with the highest productivity in sector A decide to work in sector A?

Individual i's choice for selecting sector B can be written as:

$$\log W_i^B > \log W_i^A$$

$$\log X_i^B + \log P^B > \log X_i^A + \log P^A$$

$$\log X_i^B > \log X_i^A + \log P^A - \log P^B$$

$$\log X_i^B > \phi \log X_i^B + \log P^A - \log P^B$$

$$\log X_i^B (1-\phi) > + \log P^A - \log P^B$$

$$\log X_i^B > \frac{\log P^A - \log P^B}{1-\phi}.$$

Note that the RHS of this inequality is a constant term (doesn't vary by i), such that the higher is individual i's productivity in sector B, the more likely is the individual to work in sector B. We can thus conclude that workers with the highest productivity in sector B will decide to work in sector B (as long as this inequality is satisfied for at least some of the workers, i.e., there is employment in both sectors of the economy).

An individual i works in sector A if

$$\log W_i^B \leq \log W_i^A$$

$$\log X_i^B + \log P^B \leq \log X_i^A + \log P^A$$

$$\log X_i^B \leq \log X_i^A + \log P^A - \log P^B$$

$$\frac{\log X_i^A}{\phi} \leq \log X_i^A + \log P^A - \log P^B$$

$$\log X_i^A \frac{(1-\phi)}{\phi} \leq \log P^A - \log P^B$$

$$\log X_i^A \leq \frac{\phi}{1-\phi} \left(\log P^A - \log P^B \right).$$

Note that the RHS of this inequality is also a constant term (doesn't vary by i). However, this inequality implies that the higher is individual i's productivity in sector A, the *less* likely is the individual to work in sector A. We can thus conclude that workers with the highest productivity in sector A would *not* decide to work in sector A (but rather in sector B).

(The answer to this question can also be given using an illustration, as provided in the slides for Lecture 7.)

(b) Discuss whether and if so how the results in (a) would change if $\phi > 1$.

Yes, the answers to (a) would be opposite if $\phi > 1$.
To see this, we must note that $(1-\phi)$ is now negative in the second-last inequalities in the derivations above, and there must be a sign reversal in the last inequality.
An individual i works in sector B if

$$\log X_i^B(1-\phi) > +\log P^A - \log P^B$$
$$\log X_i^B < \frac{\log P^A - \log P^B}{1-\phi}.$$

An individual i works in sector A if

$$\log X_i^A \frac{(1-\phi)}{\phi} \leq \log P^A - \log P^B$$
$$\log X_i^A \geq \frac{\phi}{1-\phi}\left(\log P^A - \log P^B\right).$$

Thus, if $\phi > 1$, workers with the highest productivity in sector A will decide to work in sector A, while workers with the highest productivity in sector B will *not* work in sector B (but rather in sector A).
(The answer to this question can also be given using an illustration, as provided in the slides for Lecture 7.)

(c) Suppose that there is an exogenous increase in the market price of goods produced in sector A. How does this affect sector selection and results in (a)?

If there is an increase in the market price of goods produced in sector A, P^A, from the choice equation for sector B we note that it is *less* likely that an individual i with a given level of sector B productivity will work in sector B:

$$\log X_i^B > \frac{\log P^A - \log P^B}{1-\phi}$$

However, it is still the case that workers with the highest productivity in sector B will decide to work in sector B as long as this inequality is satisfied for at least some of these workers.

From the choice equation for sector A, we note that it is now *more* likely that an individual i with a given level of sector A productivity will work in sector A:

$$\log X_i^A \leq \frac{\phi}{1-\phi}\left(\log P^A - \log P^B\right)$$

However, it is still the case that workers with the highest productivity in sector A would *not* decide to work in sector A as long as this inequality is not satisfied for at least some of these workers.

(The answer to this question can also be given using an illustration, as provided in the slides for Lecture 7.)

(d) Discuss how sector selection depends on the variance of log-productivities in (a) and (b). Explain the intuition behind this type of selection in the labor market.

This is a Roy model of sector selection where workers' productivities across sectors are perfectly positively correlated. When sectoral productivities are perfectly positively correlated, we know that workers with the highest (lowest) productivity in one sector will also have the highest (lowest) productivity in the other sector. *In this case, the selection of sectors depends on the variance of log-productivities across sectors, such that the more productive workers tend to select into the sector with the highest variance.* To see this, we can take the variance (across i) of the following relationship:

$$\log X_i^A = \phi \log X_i^B$$

$$var(\log X^A) = \phi^2 var(\log X^B)$$

If $\phi \in (0,1)$ as in (a) then $var(\log X^A) < var(\log X^B)$, such that there is a positive selection to sector B. However, if $\phi > 1$ as in (b) then $var(\log X^A) > var(\log X^B)$, such that is positive selection to sector A. In other words, the most productive workers decide to go to the sector that has the highest variance in log-productivity $var(\log X^s)$, and as a consequence, the highest variance in log-earnings $var(\log W^s)$ or returns to workers' skills.

3. This question is about: Aggarwal, R. K. and A. A. Samwick. (1999). The Other Side of the Trade-off: The Impact of Risk on Executive Compensation. *Journal of Political Economy* 170(1), 65-105.

(a) Explain the source and the intuition behind the fundamental trade-off in a principal-agent model.

> The source of the principal-agent problem lies in an information asymmetry: There is a risk averse agent who knows his own effort, but the principal only observes the agent's performance which is a noisy signal of effort. Thus, contracts cannot be made contingent on effort, only on performance. The firm would like to offer a wage which only depends on performance since this provides the agent with the optimal incentive to put in effort. The agent on the other hand prefers a fixed wage since a piece rate exposes him to uncertainty (the noisy part of performance) which he dislikes. The risk neutral firm can therefore insure worker by offering a wage contract that is part performance pay and part fixed wage. And since the agent is willing to pay for the insurance, the optimal contract therefore lies somewhere in between full insurance and no insurance. This illustrates the trade-off in the model between incentives and insurance. Without the asymmetric information, there would be no such trade-off: The contract could be made contingent on effort. This provides the optimal incentive and does not expose the agent to risk.

(b) First, by solving a traditional principal-agent model where agents' compensation w depends on firm performance π, i.e., $w = \alpha_0 + \alpha_1\pi$, the authors find the solution to the performance-pay related component of the optimal contract as follows:

$$\alpha_1^* = \frac{1}{1 + \gamma\sigma_\pi^2} \tag{1}$$

Note that γ is a positive constant which depends on agents' risk preferences and disutility of effort. The standard deviation of firm performance is σ_π.

Next, a relative performance evaluation scheme is considered, where agents' compensation also depends on industry performance κ, i.e., $w = \tilde{\alpha}_0 + \tilde{\alpha}_1\pi + \tilde{\alpha}_2\kappa$. Suppose $\pi = \rho\frac{\sigma_\pi}{\sigma_\kappa}\kappa + \varepsilon$, where ε is an *idiosyncratic* firm-level performance shock such that $cov(\kappa, \varepsilon) = 0$. We denote the standard deviation of industry performance by σ_κ and the correlation of firm and industry performance by ρ. The optimal performance-related components in this case are:

$$\begin{cases} \tilde{\alpha}_1^* = & \frac{1}{1+\gamma\sigma_\pi^2(1-\rho^2)} \\ \tilde{\alpha}_2^* = & -\tilde{\alpha}_1^*\rho\frac{\sigma_\pi}{\sigma_\kappa} \end{cases} \tag{2}$$

Suppose that $\rho > 0$. Explain the intuition behind $\tilde{\alpha}_1^* > \alpha_1^*$ and $\tilde{\alpha}_2^* < 0$.

Firm performance π depends on both industry performance κ (common factors across firms in the industry) and idiosyncratic firm performance ε, and moreover, industry performance κ is orthogonal to idiosyncratic firm performance ε, i.e., $cov(\kappa, \varepsilon) = 0$. Thus, one can interpret ε as a better proxy for the executive's actions (*agent's effort*) than the observed firm performance π. In this case, the shareholder (*principal*) would like to offer a performance pay scheme that depends only on the idiosyncratic firm performance ε. However, this is not feasible since the idiosyncratic component is not directly observed by the principal. If information on *both* firm performance π and industry performance κ is available, the principal can design a contract where performance pay depends on ε, i.e., $w = \tilde{\alpha}_0 + \tilde{\alpha}_1\varepsilon$. To see this, let's use $\pi = \rho\frac{\sigma_\pi}{\sigma_\kappa}\kappa + \varepsilon$ and $\tilde{\alpha}_2 = -\tilde{\alpha}_1\rho\frac{\sigma_\pi}{\sigma_\kappa}$, and after these inserting in $w = \tilde{\alpha}_0 + \tilde{\alpha}_1\pi + \tilde{\alpha}_2\kappa$, we get:

$$w = \tilde{\alpha}_0 + \tilde{\alpha}_1\pi + \tilde{\alpha}_2\kappa = \tilde{\alpha}_0 + \tilde{\alpha}_1\left(\rho\frac{\sigma_\pi}{\sigma_\kappa}\kappa + \varepsilon\right) - \tilde{\alpha}_1\rho\frac{\sigma_\pi}{\sigma_\kappa}\kappa$$

$$\Rightarrow w = \tilde{\alpha}_0 + \tilde{\alpha}_1\varepsilon$$

The intuition for having an optimal contract where $\tilde{\alpha}_2 = -\tilde{\alpha}_1\rho\frac{\sigma_\pi}{\sigma_\kappa} < 0$ is as follows: *in order to reward the firm's executive (agent) according to idiosyncratic firm performance, which is thought to proxy exerted effort and which is uncorrelated to common factors across all firms in the industry ($cov(\kappa, \varepsilon) = 0$), the principal has to set a penalty on industry performance κ, when industry performance is positively correlated ($\rho > 0$) to firm performance ("noise" part).* Next, we need to explain why the performance pay component related to firm performance is higher under a relative performance evaluation than under an absolute performance pay, i.e., $\tilde{\alpha}_1 > \alpha_1$. *The reason for this is that the optimal $\tilde{\alpha}_1$ depends on the variance of the idiosyncratic firm performance $\sigma_\pi^2(1-\rho^2)$ rather than the variance of overall firm performance σ_π^2. The intuition behind this result is that by exploiting both firm performance and industry performance, the principal can provide stronger incentives on firm performance than would be feasible if only a measure of firm performance was available.* To see this, we can take the variance of $\pi = \rho\frac{\sigma_\pi}{\sigma_\kappa}\kappa + \varepsilon$ and by using $cov(\kappa, \varepsilon) = 0$ we get:

$$\sigma_\pi^2 = \rho^2\frac{\sigma_\pi^2}{\sigma_\kappa^2}\sigma_\kappa^2 + \sigma_\varepsilon^2 \Rightarrow \sigma_\varepsilon^2 = \sigma_\pi^2(1-\rho^2)$$

From eq (2), we note that the optimal $\tilde{\alpha}_1$ depends on σ_ε^2 rather than σ_π^2.
[The students are asked to provide the intuition behind $\tilde{\alpha}_1^* > \alpha_1^*$ and $\tilde{\alpha}_2^* < 0$. Answers that include sentences marked out in italics above are thus sufficient. Derivations of these results in not required, but can be helpful.]

4. This question is about: Fehr, E. and L. Goette. (2007). Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment. *American Economic Review* 97(1), 298-317. The authors conducted a randomized field experiment at a bicycle messenger service in Zurich, Switzerland. They randomly assigned bicycle messengers working at a company called Veloblitz, who were willing to participate in the experiment, to two groups. For group A, they implemented a 25-percent increase in the commission rate during the four weeks in September 2000. The messengers in group B were paid their normal commission rate during this time period. During the four weeks in November 2000 Group B received a 25-percent increase in the commission rate, while the members of group A received their normal commission rate. Table 1 shows the main results from this paper.

Table 1. Results from Fehr and Goette (2007)

Part A

MAIN EXPERIMENTAL RESULTS
(*OLS regressions*)

	Dependent variable: Shifts per four-week period		
	(4)	(5)	(6)
Observations are restricted to	Messengers participating in experiment	All messengers at Veloblitz	All messengers at Flash and Veloblitz
Treatment dummy	3.99***	4.08***	3.44**
	(1.030)	(0.942)	(1.610)
Dummy for nontreated at Veloblitz			−0.772
			(1.520)
Treatment period 1	−1.28	−1.57	−0.74
	(1.720)	(1.210)	(0.996)
Treatment period 2	−2.56	−2.63**	−2.19**
	(1.860)	(1.260)	(1.090)
Individual fixed effects	Yes	Yes	Yes
R squared	0.694	0.74	0.695
N	124	190	386

Note: Robust standard errors, adjusted for clustering on messengers, are in parentheses.
*** Indicates significance at the 1-percent level.
** Indicates significance at the 5-percent level.
* Indicates significance at the 10-percent level.
Source: Own calculations.

Part B

THE IMPACT OF THE EXPERIMENT ON LOG REVENUES PER DAY
(Dependent variable: log (revenues per shift) during fixed shifts, OLS regressions)

	(1)	(2)
Treatment dummy	−0.0642**	−0.0601**
	(0.030)	(0.030)
Gender (female = 1)	−0.0545	
	(0.052)	
Log(tenure)	0.105***	0.015
	(0.016)	(0.062)
Day fixed effects	Yes	Yes
Individual fixed effects	No	Yes
R-Squared	0.149	0.258
N	1,137	1,137

Note: Robust standard errors, adjusted for clustering on messengers, are in parentheses.
*** Indicates significance at the 1-percent level.
** Indicates significance at the 5-percent level.
* Indicates significance at the 10-percent level.
Source: Own calculations.

(a) Interpret the result in Part A of Table 1, column (4), row (1), which shows a point estimate of 3.99. Is this result consistent with the predictions of a standard neoclassical model of intertemporal labor supply? Explain why or why not.

The intertemporal model of labor supply predicts that labor supply will be high in periods when the wage is high, and lower in periods with lower wages. The coefficient in Part A of Table 1 in column (4)-row (1) shows a positive and significant treatment effect; the treated group works on average four shifts more than the control group. This is consistent with the intertemporal model of labor supply because labor supply as measured by the number of shifts worked is indeed positively affected by the increase in the comission rate.

(b) Interpret the result in Part B of Table 1, column (2), row (1), which shows a point estimate of -0.0601. Is this result consistent with the predictions of a standard neoclassical model of intertemporal labor supply? Explain why or why not.

The intertemporal model of labor supply predicts that labor supply will be high in periods when the wage is high. The coefficient in Part B of Table 1 in column (2)-row (1) shows a negative and significant treatment effect; the wage increase leads to a reduction in revenue per shift of roughly 6 percent. This is inconsistent with the standard intertemporal model of labor supply because labor supply as measured by revenue per shift (or effort) is negatively affected by the increase in the comission rate.

(c) Is the result in Part B of Table 1, column (2), row (1), consistent with the predictions of a model with reference dependent preferences? Can an alternative neoclassical labor supply model explain this result? Explain why or why not.

Yes, the result in part B of Table 1 in column (2)-row (1) is consistent with the predictions of a model with reference dependent preferences. Suppose reference dependent behavior is reflected by the following one-period utility function:

$$U(e_t) = \begin{cases} \lambda\left(w_t e_t - \widetilde{y}\right) - g(e_t, x_t) & if\ w_t e_t \geq \widetilde{y} \\ \gamma\lambda\left(w_t e_t - \widetilde{y}\right) - g(e_t, x_t) & if\ w_t e_t < \widetilde{y} \end{cases}$$

where $\gamma > 1$ measures the degree of loss aversion and $\widetilde{y}$ is the daily income target. The increase in wages makes it more likely that the income target is already met or exceeded at relatively low levels of effort for individuals in the treatment group. Therefore, compared to the control group, the workers in the treatment group are more likely to face a situation where the marginal utility of income is λ instead of $\gamma\lambda$, i.e., they face lower incentives to work during the shift. As a consequence, members of the treatment group will provide less effort than members of the control group during a day (shift).

Yes, this result can also be explain by a neoclassical model with nonseparable utility where the disutility of effort in period t depends on the effort exerted in period t-1, as reflected by the following one-period utility function:

$$U(e_t, e_{t-1}) = \lambda w_t e_t - g(e_t(1 + \alpha e_{t-1}), x_t)$$

When $\alpha > 0$, the disutility of exerting additional effort in period t increase as a function of effort exerted in period *t-1*, as follows:

$$\frac{dg(e_t(1 + \alpha e_{t-1}), x_t)}{de_t} = g_1^{'}(e_t(1 + \alpha e_{t-1}), x_t)(1 + \alpha e_{t-1})$$

Thus, messengers in the treatment group who work more shifts when the wage is high may rationally decide to reduce the effort per shift because exerting effort across all shifts is more costly. However, the model above does not predict that workers who work more shifts (days) will *necessarily* reduce their effort per shift. It simply allows for this possibility. If the wage increase is large enough, it is also possible that workers who behave according to this model raise their effort per shift.

5. This question is about: Ewens, M., B. Tomlin, and L. C. Wang. (2014). Statistical Discrimination or Prejudice? A Large Sample Field Experiment. *Review of Economics and Statistics* 96(1), 119-134.

   (a) Explain the sources of differential treatment of socio-economic groups in models of taste-based discrimination and models of statistical discrimination.

   Let's consider a case with two equally productive individuals that belong to different socio-economic groups. At the group level, average productivities across the two socio-economic groups vary, even though the two individuals we consider are assumed to have identical productivity.
   Sources of differential treatment in a model of taste-based discrimination:

   - the evaluator observes productivities of both individuals, yet discriminates against one of individuals, because the evaluator *holds a preference or a taste for discrimination* against one of the socio-economic groups.

   Sources of differential treatment in a model of statistical discrimination:

   - the evaluator doesn't observe individual productivities, yet knows differences in average productivities across groups, and uses group differences to predict individual productivities and therefore treats equally productive individuals differently on the basis of this prediction

   - if, in addition, the evaluator receives (noisy) signals of individual productivities, differences in how the evaluator weights signals received from individuals belonging to different groups may also lead to differential treatment of equally productive individuals

   - thus, *differences in group averages* and *differences in the weighting of signals from different groups* may lead to statistical discrimination when individual productivity is not observed by the evaluator.

   (b) Consider a situation where applicants (e.g., job-seekers, renters, etc.) from two socio-economic groups, majority A and minority B, send applications to an evaluator (e.g., employer, landlord, etc.). Applicants can either send a positive signal, a negative signal, or no signal. The evaluator is from the majority group A. The average quality of group A applicants in the population is higher. Discuss (or illustrate) group differences in interview probabilities under the following two cases: i) the majority evaluator has no taste for discrimination, but predicts applicant quality based on received signals, and ii) the majority evaluator has a taste for discrimination against group B (out-group discrimination).

To answer (b), case i), the students can draw and explain a figure similar to Figure 1 (Case 2: $\gamma_W > \gamma_B$) on page 123 in Ewens et al. (ReStat 2014).
Since majority applicants are on average of a higher quality, their 'interview probability curve' lies above the curve for minority applicants. Moreover, since the evaluator belongs to the majority group, the slope of the curve for majority applicants is higher as their signals are weighted more by evaluators. The following explanations should be provided:

- If no signal is provided, the evaluator will use average group differences to predict the quality of applicants, and the interview probability for applicants from the majority group will be higher (baseline gap)

- If a negative signal x^- (below minority average) is given, interview probabilities for applicants from both groups would decline. However, the decline is higher for the majority applicants, since i) by giving a negative signal they move far away from their no-signal mean, and ii) their signals are also weighted higher by the evaluator. As a consequence, the differences in interview probabilities across majority and minority applicants would be smaller when a negative signal is given relative to the (no-signal) baseline gap.

- If a positive signal x^+ (above majority average) is given, interview probabilities for applicants from both groups would increase. However, the increase is higher for the minority applicants, since they now indicate having a much higher quality relative to their no-signal mean. Nonetheless, since the majority applicants also experience an increase in their interview probabilities, it is unclear whether differences in interview probabilities across majority and minority applicants would be smaller or larger when a positive signal is given relative to the baseline gap.

To answer (b), case ii), the students should be able to argue that a similar relationship between interview probabilities, signal quality and group identities as illustrated/discussed above in a) would exist if the majority evaluator exhibits out-group discrimination.

We can introduce out-group discrimination by adding a coefficient of discrimination $k \in (0, 1)$ that enters negatively in the evaluator's expected utility of interviewing a minority applicant, as follows:

$$E[U(\theta_i)] = (1 - k \cdot \mathbf{1}(g = B)) E[\theta_i]$$

where θ_i is the quality of applicant i.

For simplicity, let's assume the evaluator's forecast of applicant quality doesn't depend on group identity, so that $\hat{\theta}_i = \mu + \gamma x_i$ (in contrast, with statistical discrimination, we allowed μ and γ to depend on the applicant's group identity). The evaluator's expected utility from group g is then given as follows:

$$E[U(\theta_i)] = (1 - k \cdot \mathbf{1}(g = B)) E[\theta_i]$$

$$\Rightarrow E[U(\theta_i)] = \begin{cases} \mu + \gamma x & if\ g = A, \\ \mu(1-k) + \gamma(1-k)x & if\ g = B. \end{cases}$$

The introduction of a coefficient of discrimination $k \in (0, 1)$ implies that the majority evaluator gets a lower expected utility from interviewing a minority applicant than a majority applicant, even when both provide the same signal x, since $\mu(1-k) < \mu$ and $\gamma(1-k) < \gamma$. This shifts the 'interview probability curve' for minority applicants below the curve for majority applicants. Moreover, $\gamma(1-k) < \gamma$ also implies that the majority evaluator receives a lower increase in expected utility from a marginal increase in signal quality for the minority applicant compared to a similar increase for the majority applicant. Thus, the slope of the 'interview probability curve' for minority applicants is lower than the slope for majority applicants. Thus, we get a similar relationship between interview probabilities for minority and majority applicants with out-group discrimination as we did in the answer to (b), case i).

(c) Consider a dataset consisting only of majority (group A) evaluators, where the researcher has randomly assigned applications across evaluators. Is it possible for the researcher to distinguish between taste-based and statistical discrimination? What if data on minority (group B) evaluators is also available? How does this depend on minority evaluators having out-group discrimination?

If the dataset consists only of majority evaluators, it will be not be possible for the researcher to distinguish between statistical discrimination and taste-based discrimination. As shown in the solution to (b) above, the relationship between interview probabilities, signal quality and group identities is similar under these two types of discrimination. The intuition for this result is that there is no variation with respect to tastes for discrimination or statistical discrimination (across evaluators) in this dataset that can allow the researcher to separately identity these two types of discrimination.

If observations on minority evaluators are also included in the dataset and these evaluators are assumed to exhibit a similar out-group discrimination with a coefficient of discrimination k (against majority group A)*, then it might be possible to separately identity the two types of discrimination.* The minority evaluators' expected utility from interviewing an applicant from group g is:

$$E\left[U(\theta_i)\right] = (1 - k \cdot \mathbf{1}(g = A))\, E\left[\theta_i\right]$$

$$\Rightarrow E\left[U(\theta_i)\right] = \begin{cases} \mu + \gamma x & if\ g = B, \\ \mu(1-k) + \gamma(1-k)x & if\ g = A. \end{cases}$$

If minority evaluators have out-group discrimination, the 'interview probability curve' for minority applicants lies above the curve for majority applicants and the slope of this curve is also higher for minority applicants. (This is actually opposite to the case we discussed in (b) above.)

In contrast, if minority evaluators don't exhibit out-group discrimination but instead statistically discriminate, then (i) the slope of 'interview probability curve' for minority applicants is higher, since minority evaluators put more weight on signals given by minority applicants, but as earlier, (ii) since majority applicants on average are of a higher quality, the 'interview probability curve' for majority applicants still lies above the curve for minority applicants.

Importantly, the effects of giving positive signals on gaps in majority-minority interview probabilities will differ in the two cases. With statistical discrimination, positive signals sent to minority evaluators will reduce the gap relative to the baseline gap. In contrast, with taste-based discrimination, positive signals sent to minority evaluators will increase the gap relative to the baseline gap. Using these differences in the predictions of statistical discrimination and taste-based discrimination for minority evaluators, it is possible to separately identity these two types of discrimination for the researchers.

If minority evaluators do not exhibit out-group discrimination then once again there is not sufficient variation in tastes for discrimination in this dataset to be able to identify the two types of discrimination.