14.03 Fall 2004
Problem set 1 Solutions

Professor: David Autor

September 29, 2004

1 Math tools - Part I

Consider the following function:

$$y = f(x; a) = a^2 + x - ax^2$$

where $a > 0$ is a parameter.

1. Find the first order condition for a critical point of this function.

$$\begin{aligned} \frac{dy}{dx} &= 0 \\ 1 - 2ax &= 0 \end{aligned}$$

2. Is this a maximum or a minimum or an inflection point?

   To determine whether this is a maximum or minimum we check the sign of the second derivative:

$$\frac{d^2y}{(dx)^2} = -2a < 0$$

   This function is concave therefore the FOC finds a maximum of this function.

3. Solve for $x^*(a)$, the maximizer of the function $f(x; a)$ .Also find $y^*(a)$, the maximized value of y as a function of a

$$\begin{aligned} x^*(a) &= \frac{1}{2a} \\ y^*(a) &= a^2 + \frac{1}{4a} \end{aligned}$$

4. Find $\frac{dx^*}{da}$ and $\frac{dy^*}{da}$

$$\begin{aligned}\frac{dx^*}{da} &= -\frac{1}{2a^2} \\ \frac{dy^*}{da} &= 2a - \frac{1}{4a^2}\end{aligned}$$

5. Now use the FOC and the implicit function theorem to find $\frac{dx^*}{da}$

$$\begin{aligned}1 - 2ax^* &= 0 \\ -2x^* da - 2a dx^* &= 0 \\ \frac{dx^*}{da} &= -\frac{x^*}{a} = -\frac{1}{2a^2}\end{aligned}$$

6. Use the envelope theorem to find $\frac{dy^*}{da}$. Why does this theorem allow you to simplify your calculations with respect to point 4?

$$\begin{aligned}\frac{dy^*}{da} &= \frac{\partial y\left(x^*\left(a\right)\right)}{\partial x^*}\frac{dx^*\left(a\right)}{da} + \frac{\partial y\left(x^*\left(a\right)\right)}{\partial a} = \\ &= 0 + 2a - \left(x^*\right)^2 = \\ &= 2a - \frac{1}{4a^2}\end{aligned}$$

The envelope theorem allows us to disregard the impact of a change in a on the maximizer x^* since at the optimum $\frac{\partial y}{\partial x^*} = 0$. With respect to part 4 we can avoid substituting $x^*\left(a\right)$ to find $y^*\left(a\right)$ and its derivative.

2 Math tools - Part II

Consider the problem

$$\begin{aligned}\max\ z &= 2x + 4x^{\frac{1}{2}}y^{\frac{1}{2}} + 2y \\ s.t.\ \ x + y &= 400\end{aligned}$$

1. Set up the lagrangian and find the first order conditions

$$\begin{aligned}L &= 2x + 4x^{\frac{1}{2}}y^{\frac{1}{2}} + 2y - \lambda\left(x + y - 400\right) \\ \frac{\partial L}{\partial x} &= 2 + 2\left(\frac{y}{x}\right)^{\frac{1}{2}} - \lambda = 0 \\ \frac{\partial L}{\partial y} &= 2 + 2\left(\frac{x}{y}\right)^{\frac{1}{2}} - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= 400 - x - y = 0\end{aligned}$$

2. Solve explicitly for x^*, y^*, z^* and λ, the Lagrange multiplier.

One trick is to notice that the $\frac{\partial L}{\partial x} = 0$ and $\frac{\partial L}{\partial y} = 0$ FOC tell us

$2 + 2(\frac{y}{x})^{\frac{1}{2}} = \lambda = 2 + 2(\frac{x}{y})^{\frac{1}{2}}$

which implies that $x = y$. Making this substitution into all of the FOC quickly gives us

$$\begin{aligned} x^* &= 200 \\ y^* &= 200 \\ z^* &= 1600 \\ \lambda &= 4 \end{aligned}$$

3. Set up the dual problem of 1., i.e., minimize the primal constraint function subject to the primal objective function being equal to the value of z^* obtained in 2.

$$\min \; x + y$$
$$s.t. \; 2x + 4x^{\frac{1}{2}}y^{\frac{1}{2}} + 2y \;=\; 1600$$

$$L = x + y - \lambda^D \left(2x + 4x^{\frac{1}{2}}y^{\frac{1}{2}} + 2y - 1600\right)$$

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - \lambda^D \left(2 + 2\left(\frac{y}{x}\right)^{\frac{1}{2}}\right) = 0 \\ \frac{\partial L}{\partial y} &= 1 - \lambda^D \left(2 + 2\left(\frac{x}{y}\right)^{\frac{1}{2}}\right) = 0 \\ \frac{\partial L}{\partial \lambda} &= -\left(2x + 4x^{\frac{1}{2}}y^{\frac{1}{2}} + 2y - 1600\right) = 0 \end{aligned}$$

4. Solve explicitly for x, y, and λ^D, the Lagrange multiplier in the dual problem.

Again the $\frac{\partial L}{\partial x} = 0$ and $\frac{\partial L}{\partial y} = 0$ FOC tell us that $x = y$. (Unless $\lambda^D = 0$.But then the $\frac{\partial L}{\partial x} = 0$ and $\frac{\partial L}{\partial y} = 0$ FOC would be saying $1 = 0$, so we know that $\lambda^D \neq 0$.) This allows us to calculate that

$$\begin{aligned} x^* &= 200 \\ y^* &= 200 \\ \lambda^D &= \frac{1}{4} \end{aligned}$$

5. What is the relationship between λ and λ^D, and can you provide an economic explanation for this?

$$\lambda^D = \frac{1}{\lambda}$$

We know that λ describes the increase in z^* associated with relaxing the constraint in the primal problem by one unit. We also know that λ^D represents the gain in the minimization problem (dual) from reducing the constraint on z by 1 unit. The relationship between the two multipliers is explained as follows: in order to increase z^* by λ units, we have to relax the constraint in the primal by one. So if we want to increase z^* by one we have to relax the constraint by $\frac{1}{\lambda}$, which is also the gain in the minimization problem by relaxing the constraint on z by one unit, that is λ^D.

3 Math tools - Part III

Let

$$U = xy^2$$

1. Derive the indirect utility function as a function of p_x, p_y and M, where p_x and p_y are respectively the prices of the two goods x and y, and where M is the consumer's income.

$$\begin{aligned} L &= xy^2 - \lambda\left(p_x x + p_y y - M\right) \\ \frac{\partial L}{\partial x} &= y^2 - \lambda p_x = 0 \\ \frac{\partial L}{\partial y} &= 2xy - \lambda p_y = 0 \\ \frac{\partial L}{\partial \lambda} &= -\left(p_x x + p_y y - M\right) = 0 \end{aligned}$$

We can use the first two FOC to find

$$y^* = \frac{2p_x}{p_y} x^*$$

Substituting this into the last FOC then allows us to calculate

$$\begin{aligned} x^* &= \frac{M}{3p_x} \\ y^* &= \frac{2M}{3p_y} \end{aligned}$$

And

$$V(p_x, p_y, M) = x^* y^{*2} = \frac{M}{3p_x} \left(\frac{2M}{3p_y} \right)^2$$

2. Now calculate the level consumption of both goods and the level of utility achieved by this consumer if prices and income are as follows:

$$p_x = 2,\ p_y = 3\,, M = 9$$
$$x^* = \frac{3}{2},\ y^* = 2$$
$$U^* = 6$$

3. Now set up the dual of this problem: minimize expenditure subject to the level of utility that you calculated in part 2 (U^*) and prices p_x and p_y. Find the expression for the expenditure function $E(p_x, p_y, U)$.

$$\begin{aligned} \min \quad & p_x x + p_y y \\ s.t.\ & xy^2 = 6 \end{aligned} \tag{1}$$

$$\begin{aligned} L &= p_x x + p_y y - \lambda^D \left(xy^2 - U \right) \\ \frac{\partial L}{\partial x} &= p_x - \lambda^D y^2 = 0 \\ \frac{\partial L}{\partial y} &= p_y - \lambda^D \cdot 2xy = 0 \\ \frac{\partial L}{\partial \lambda} &= -\left(xy^2 - U \right) = 0 \end{aligned}$$

You can use the first two FOC to find that $y^* = 2\frac{p_x}{p_y}x^*$. Then substitute into the last FOC to get

$$\begin{aligned} x^* &= \left(\frac{U}{4}\right)^{\frac{1}{3}} \left(\frac{p_y}{p_x}\right)^{\frac{2}{3}} \\ y^* &= 2\left(\frac{U}{4}\right)^{\frac{1}{3}} \left(\frac{p_x}{p_y}\right)^{\frac{1}{3}} \\ E\left(p_x, p_y, U\right) &= 3\left(\frac{U}{4}\right)^{\frac{1}{3}} p_x^{\frac{1}{3}} p_y^{\frac{2}{3}} \end{aligned}$$

You can verify that:

$$E\left(2, 3, 6\right) = 9$$

4 Labor market structure

a. In a competitive market the firm chooses the amount of labor L given a wage w and a price p in order to maximize profits:

1. (a)

$$\begin{aligned} \max_L \Pi &= p(-0.5L^2 + 10L) - wL \\ \frac{\partial \Pi}{\partial L} &= 0 \Rightarrow w = -2L + 20 \end{aligned}$$

The above equation identifies a labor demand curve.
In order to find market equilibrium we will have to equate demand and supply for labor:

$$\begin{aligned} L^D &= 10 - 0.5w \\ L^S &= -10 + w \\ L^D &= L^S \end{aligned}$$

The equilibrium will be given by:
$L^C = 3.33, \quad w^C = 13.33$

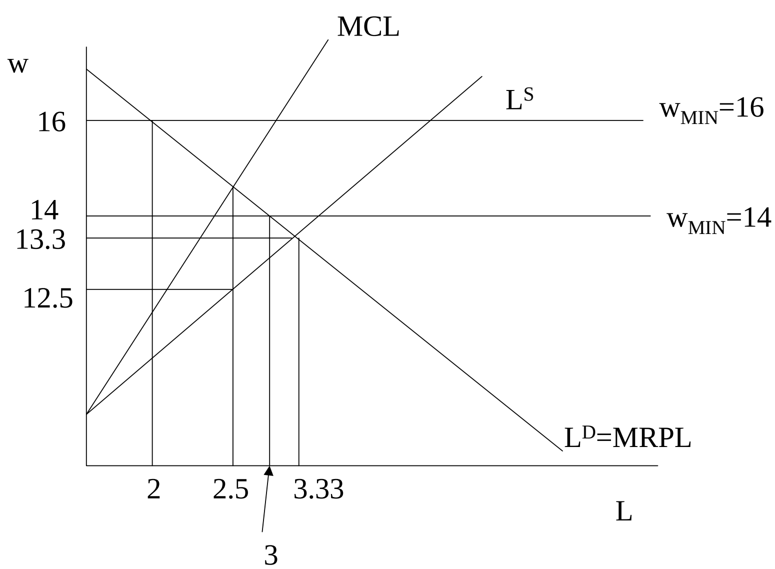
b. A monopsonist faces an upward sloping supply of labor therefore the marginal cost of labor faced by the monopsonist is increasing.

The monopsonist chooses the quantity of labor in order to maximize profits. He will equate Marginal Revenue Product of Labor (MRPL) to the Marginal Cost of Labor (MCL):

$$\max_L pY - w(L)L$$
$$(MRPL) \quad p\frac{\partial Y}{\partial L} = w(L) + \frac{\partial w(L)}{\partial L}L \quad (MCL)$$
$$2(-L+10) = 10 + 2L$$

Therefore we will observe an equilibrium:

$$L^M = 2.5, \quad w^M = 12.5$$



c. See Graph.

d. The monopsonistic employment level is lower and and the wage is lower than the MRPL, which gives the monopsonist positive rents. The reason this happens is the following: The single firm in a competitive market faces a flat supply for labor while the monopsonist faces an upward sloping labor supply. The monopsonist takes into account the fact that when he wants to hire one more worker, he will have to pay a higher wage to this worker and to all the other workers. This effect induces the profit maximizing monopsonist to restrict employment with respect to the competitive market case.

e. We expect the employment level to move in opposite direction according to the market structure.

In a competitive market the employment level would fall:

$$\begin{aligned} w^C_{MIN} &= 14 \\ L^C_{MIN} &= 3 \end{aligned}$$

In a monopsonistic market at this wage level the employment would increase:

$$\begin{aligned} w^M_{MIN} &= 14 \\ L^M_{MIN} &= 3 \end{aligned}$$

This last result is due to the fact that the minimum wage makes the labor supply flat over some range and this eliminates the degree of market power that makes the monopsonist inefficiently keep a low level of employment. Notice that equilibrium wages and quantities are the same here regardless of market structure–the minimum wage transforms the monopsonist into a price taker.

f. If the minimum wage were fixed at this higher level we would observe a decline in employment in both kinds of market, so this kind of experiment would be uninformative as to whether the labor market has a competitive structure or not.

$$\begin{aligned} w^C_{MIN} &= 16 \\ L^C_{MIN} &= 2 \end{aligned}$$

$$\begin{aligned} w^M_{MIN} &= 16 \\ L^M_{MIN} &= 2 \end{aligned}$$

5 RCM

You are conducting an experiment to assess the causal effect of taking low doses of aspirin (LDA) on blood pressure. Consider each of the following experimental approaches. For each, what are the assumptions necessary to draw a causal inference? How plausible are these assumptions? In aswering the following questions make use of the setup developed in the lecture on Causal Inference where T^* is the true effect of the treatment and Y is the outcome of interest, blood pressure. Y can take up two values: $Y = 0$ if blood pressure is normal and $Y = 1$ if blood pressure is high. We indicate the treatment by X, which can take two values: $X = 1$ if LDA are given to the individual and $X = 0$ if no aspirin is given. Also indicate time by $t = 0$ or $t = 1$. The true treatment effect on the individual i is the following:

$$T_i^* = Y_{i1,t} - Y_{i0,t}$$

For each of the following scenarios, please describe the assumptions required to obtain a valid estimate of T_{it}^* or $E\left[T_{it}^*\right]$ (the latter denotes the average treatment effect for the treated population). Describe for each setup how your estimate of T^* might be biased if your assumptions are not satisfied.

A couple of general point about this question: Many people mentioned small sample sizes as a source of bias. Usually we call an estimator biased only if its expected value differs from the parameter we're trying to estimate. Small sample sizes will make for a noisy estimator, but we would not generally say that they create bias since on average our estimator will still be right (barring other problems).

There was also considerable misunderstanding of the term "unit homogeneity". Unit homogeneity means that every individual in the sample would have the same outcome without treatment and would have the same outcome with treatment. It does not simply mean that the treatment and control groups are identical on average.

1. You choose 2,000 people at random from the entire U.S. population. You give LDA treatments to 1,000 patients and prevent the other 1,000 from using aspirin. At the end of one year, you compare the blood pressure level of treated to control patients (over the year) to measure the causal effect of LDA.

In order to estimate $E\left[T_{it}^{*}\right]$, we need to assume that the patients assigned LDA were selected randomly; more precisely, we assume that $E\left[T_{it}^{*}|X=0\right] = E\left[T_{it}^{*}|X=1\right]$ (the expected treatment effect is the same for the two groups). If we instead want to know T_{it}^{*} – the treatment effect for a specific individual – we must assume unit homogeneity: all units are the same. In that case, observing two units, one exposed to the treatment and one not exposed, is the same as observing the same unit exposed to the treatment and not exposed to the treatment. With unit homogeneity,

$$Y_{i1} = Y_{j1}$$

and therefore the treatment effect can be measured as

$$T_i = Y_{j1} - Y_{i0}$$

Unit homogeniety is almost certainly false here, but random assignment is done all the time. Suppose, however, that some people are assigned LDA because they have high cholestorol (perhaps their doctor feels sorry for them and fudges the random assignment).This introduces a bias toward finding that LDA increases cholestorol, and we can no longer make a valid causal inference.

Notice that in this experiment we're measuring only the post-experiment blood pressures of the two groups. Many people seemed to think this was a difference-in-defference setup; here we take only one difference (across groups, but not acros time).

2. You choose 1,000 people at random from the entire U.S. population. In the 1st year, you prevent aspirin use and measure the level of blood pressure in this population. In the 2nd year, you give LDA and measure the level of blood pressure in this population. You compare the heart attack rate in the 1st and 2nd years to measure the causal effect.

   Here we assume temporal stability and causal transience: we assume that treatment and non-treatment outcomes are constant and that failing to give LDA in the 1st year does not affect cholestorol levels under LDA in the 2nd year. (As some people pointed out, causal transience also means that any aspirin people have taken in the past does not continue to have effects during the experiment.)

In math, we assume that $Y_{i1,1} = Y_{i1,0}$ and that $Y_{i0,1} = Y_{i0,0}$. If these asumptions fail then the estimate of T depends on the order in which the treatment is given, but generally T is biased. For example, suppose a health craze hits between the 1st and 2nd years; cholestorol levels improve, but this has nothing to do with LDA.

Note that there is no point in estimating $E[T_{it}^*]$ here. Once we make the above assumptions, we can estimate the causal effect for each individual separately.

3. You choose 2,000 people at random from the entire U.S. population. In the 1st year, you prevent aspirin use altogether and measure the level of blood pressure. In the 2nd year, you gave LDA to 1,000 members of this population (selected at random) and continue to prevent aspirin use for the other 1,000. To measure the causal effect of LDA, you contrast the 1st to 2nd year change in the level of blood pressure in the LDA-treated population to the 1st to 2nd year change in the level of blood pressure in the non-aspirin population.

   This is the difference-in-difference approach that we saw in class. Randomization helps us build the counterfactual that we cannot observe and in doing this it allows causal inference. We adjust for any random heterogeneity between the two groups by measuring the level of blood pressure before the treatment:

	Before	After	Δ
Treat	$Y_{0i,0}$	$Y_{1i,1}$	ΔY_i
Control	$Y_{0j,0}$	$Y_{0j,1}$	ΔY_j

   $Y_{0i,0}$ is blood pressure in the treatment group before the treatment, $Y_{1i,1}$ is the blood pressure in the treatment group after the treatment, $Y_{0j,0}$ is blood pressure in the control group before the treatment to the first group and $Y_{0j,1}$ is blood pressure after the treatment to the first group. The estimated treatment effect is $T = \Delta Y_i - \Delta Y_j$.

   The question tells us that assignment was random, but without random assignment we could have biases in our estimate. For example, suppose that the people assigned to take LDA tend to be people whose cholestorol levels are rising; this would bias our estimator toward finding that LDA increases cholestorol.

2. You recruit 10 pairs of identical twins for your study. 1 of each pair receives LDA

and the other of each pair is prevented from taking aspirin. After one year, you compare the level of blood pressure in these two populations to measure the causal effect.

If twins are identical then we are in the unit homogeneity case and we just need to look at the difference between blood pressure levels in the two twins after the treatment: $T = Y_{1i,1} - Y_{0i,1}$.If unit homogeneity fails (because even identical twins have been exposed to different lifetime stresses), then we cannot estimate T^*_{it}. We could still estimate $E\left[T^*_{it}\right]$ if we assigned LDA to one twin from each pair randomly.

3. You recruit 10 pairs of fraternal twins for your study, where one of each pair is female and the other is male. The male in each pair receives LDA and the female in each pair is prevented from taking aspirin. After one year, you compare the level of blood pressure between males and females to measure the causal effect.

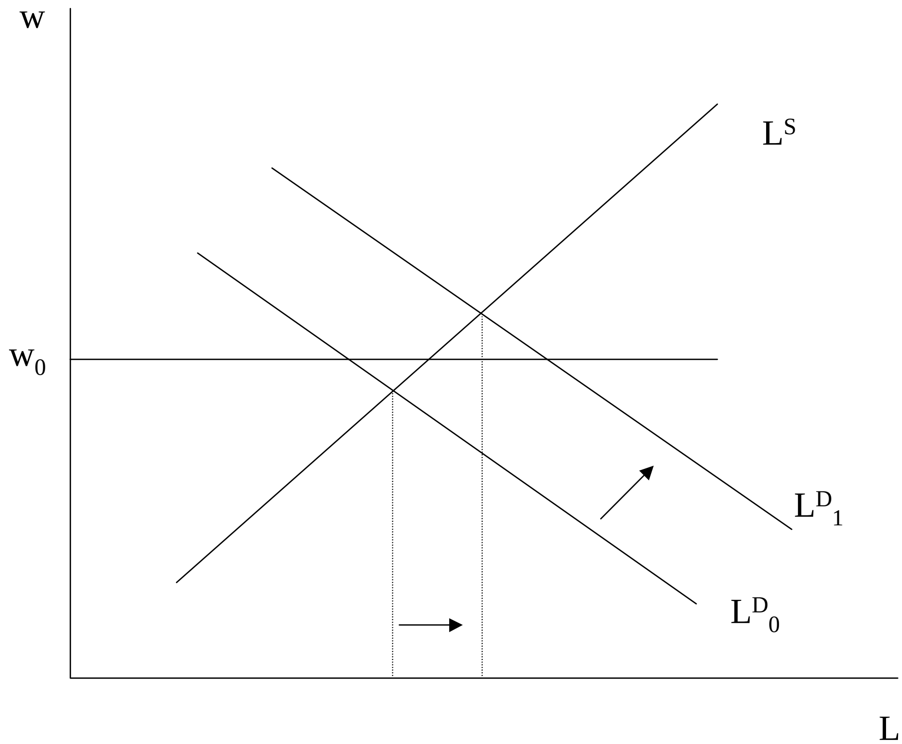
   We need to assume unit homogeneity to make a valid inference here: despite the gender and genetic differences, the treatment effect is the same for each twin in a pair. If gender has a systematic effect on the level of blood pressure, then we might go back to the difference-in-difference approach in order to recover the treatment effect T. Even here, it is possible that men and women have different trends in cholestorol levels (say men's cholestorol increases faster with age), which would invalidate the difference-in-difference strategy.

6 Card and Krueger

1. Card and Krueger (p. 773) emphasize that one strength of their analysis is that the implementation of the New Jersey minimum wage occurred during a recession.

   (a) Consider the case where the minimum wage was implemented right before an economic expansion instead. Draw a diagram that clearly shows how an economic expansion affecting both New Jersey and Pennsylvania might have invalidated the results of the 'natural experiment.'

   If the minimum wage increase had taken place during a period of fast expansion labor demand might increase so much that the new equilibrium wage is above the new minimum wage level. In the graph below say NJ and PA had the same labor demand L_0^D before the change in legislation. Then labor demand expands to L_1^D in both states, but NJ imposes a minimum wage w_0. In both states we observe an increase in the employment level, but, absent the expansion we would have observed a decrease in employment.



(b) Someone points out to you that, according to Table 3 of C&K, fast food employment fell in PA but it barely budged in NJ. Hence, they say, if we just look at NJ, it seems pretty clear that there is no support for a monopsony scenario. How would you respond to this criticism within the RCM framework? Draw a carefully labeled diagram that illustrates your point.

If NJ and PA are in fact very similar economies the decline in employment in PA would have been experienced by NJ as well, absent the minimum wage legislation change. Therefore we can conclude that according to this experiment the net effect of the new minimum wage legislation is to increase employment. We can see this using the difference-in-difference approach developed in class. Indicate Y as the employment level: $Y_{Xi,t}$ is the level of employment for store i when treatment status is X ($X = 0$ if no change in minimum wage legislation,

$X = 1$ is change in legislation), at time t.

	Before	After	Δ
Treat	$Y_{0i,0}$	$Y_{1i,1}$	ΔY_i
Control	$Y_{0j,0}$	$Y_{0j,1}$	ΔY_j

The case described above points to $\Delta Y_i \simeq 0$ and $\Delta Y_j < 0$ so the time effect $\Delta < 0$ but the treatment effect $T = \Delta Y_i - \Delta Y_j > 0$.

2. Richard Freeman distinguishes between the short and long run impacts of minimum wage increases, implying that the long run impacts are likely to be larger (and more negative). List one or two specific economic reasons why the long run response would be larger. Consider that labor is not the only factor in fast food restaurants' production functions.

   We can plausibly think that the production function allows a higher degree of substitutability between factors in the long run than in the short run. In the short run, some factors of production might be difficult to adjust, therefore the demand for labor might be more rigid. In the long run as a response to an increase in the price of one factor (labor in this case) factor proportions might vary in favor of relatively cheaper inputs and away from the relatively more expensive labor input. For example, in the short run, fast food restaurants have to keep approximately the same number of workers just to keep up with demand. In the long run, they can potentially develop and install more highly automated equipment, such as grills and cash registers, to reduce the number of workers per million burgers served. It might not have been cost-effective to do this when the wage was \$4.25 (this equipment is expensive) but at \$5.05 per hour, some additional automation will be worth doing – but this takes time.