Lecture Note: Social (Non-Market) Interactions

David Autor
MIT 14.663 Spring 2009

May 12, 2009

1 Non-Market Interactions

Non-market interactions are interactions among individuals that are not regulated by the price mechanism. They are a particular form of externality. Schelling is the intellectual father of this branch of research. It has come some ways since Schelling, but perhaps not as far as one would have anticipated.

1.1 The 'social multiplier'

The social multiplier is a well-worn term, so you should be familiar with it. Let S equal the level of social capital. We don't need to define what social capital is for now; we are only concerned with what it does. Assume for the moment that everyone takes S as given ($S = S_0$). Individuals have the utility function:

$$U = U\left(x, y; S\right).$$

Let y be the numeraire good and I be income. Agents maximize utility subject to

$$I = p_x x + y,$$

leading to the standard condition:

$$\frac{U_x}{U_y} = p_x.$$

How does a rise in S affect demand for x? It depends on the relative complementarity of S to y and x: U_{xS}, U_{yS}. So, rewriting the above condition:

$$U_x = p_x U_y.$$

If $U_{xS} > p_x U_{yS}$, then consumption of x will increase and consumption of y will fall with a rise in S. So,

$$\text{Sign}\left\langle \frac{dx}{dS} \right\rangle = \text{Sign}\left\langle U_{xS} - p_x U_{yS} \right\rangle.$$

Notice that this model does not assume that an increase in 'social capital' increases utility; it's possible that $U_S < 0$. What matters is how social capital changes the relative value of other economic choices. For example if S is drug consumption by one's peers, x is own drug consumption, and y is all other goods, it's plausible that a rise in S would raise own consumption of x and lower utility.

Let's assume that social capital of the peer group is simply equal to the average x of its members:

$$S = X = \frac{1}{N} \sum_j x^j.$$

Assume that N is large so no individual significantly affects S. Each member of the group maximizes utility by choosing x taking S as a given.

The demand function for each individual is:

$$x^j = d^j \left(e^j, p, S = X \right), \text{ where } j = 1, 2, ..., N.$$

e^j is an idiosyncratic taste variable and X is the level of social capital *assumed* by j when choosing x^j. So,

$$X = \frac{\sum d^j \left(e^j, p, X \right)}{N} = \frac{\sum x^j}{N}, \text{ or } X = F\left(e^1, e^2, ..., e^n, p \right).$$

If S and x are strong complements, then person-level idiosyncratic taste shocks have little effect on person level-demand because S cannot be adjusted simultaneously by the individual. Thus, each individual is heavily influenced by social capital.

However, if a shock hits all participants simultaneously, the cumulative effect of individual choices on social capital may matter greatly. One such shock is a change in p:

$$\frac{dS}{dp} = \frac{dX}{dp} = \frac{\sum dx^j/dp}{N} = \frac{\sum \partial x^j/\partial p}{N} + \frac{\sum \left(\partial x^j/\partial S \right) \times \left(dS/dp \right)}{N}.$$

Rearranging:

$$\begin{aligned} \frac{dS}{dp} \left(1 - \frac{1}{N} \sum \left(\partial x^j/\partial S \right) \right) &= \frac{1}{N} \sum \partial x^j/\partial p, \\ \frac{dS}{dp} &= \frac{\frac{1}{N} \sum \partial x^j/\partial p}{1 - m}, \text{ where } m = (1/N) \sum \partial x^j/\partial S. \end{aligned}$$

Here m is the famed social multiplier. The larger is m, the more a common shock to individual demands raises aggregate demand by *more* than could be predicted exclusively by summing individual demand functions.

This is kind of a nice insight: an equilibrium that is locally stable, with well-behaved demand curves at the individual level, can behave in a very non-standard fashion if a shock (such as a price drop) causes a simultaneous change in demands of many participants at once. Even though each participant has standard preferences holding other participants' behavior constant, the spillover from the consumption of each consumer to the preference of others can make the system highly reactive to common shocks.

1.2 Becker-Murphy fad model (see also Becker restaurant paper, 1991)

Write the aggregate demand for a good as:

$$Q = D\left(p, Z; Q\right), \text{ with } D_p < 0 \text{ and } D_Q > 0,$$

and Z are other demand shifters. Holding all else constant, demand is decreasing in price. But the assumption that $D_Q > 0$ means that demand for the good increases when its popularity rises (i.e., more people are consuming it).

Using the notation above,

$$\frac{dQ}{dp} = \frac{D_p}{1-m}.$$

The sign of this is ambiguous. Assume there exists a Q^* where $m = 1$ for $Q = Q^*$, $m > 1$ for $Q > Q^*$ and $m < 1$ for $Q > Q^*$. This gives rise an unstable demand curve. Demand is upward sloping below Q^*, downward sloping above Q^*, and reaches an inflection point at Q^*.

See Figure 9.1 of Becker-Murphy. If the 'market-clearing' price is on the upward sloping section of the demand curve ($Q < Q^*$), the firm can raise p further without incurring a loss in demand (in fact, just the opposite). If output is capped, then it will clearly make sense to raise p to increase quantity towards Q^*. This will create excess demand, but the profit earned per each unit sold will be higher (assuming that the rationing process is not costly). If the monopolist raises the price all the way to where $Q = Q^*$, demand becomes extremely unstable. A tiny shock in either direction could cause demand to fall precipitously. Thus, it might be profit-maximizing to a choose p so that Q is slightly less than Q^*. (This is one possible explanation for why fads—like popular restaurants—are so fleeting.)

In interpreting cases where $m > 1$, it is important to observe that the positive slope of the demand function does *not* mean that individual level demand rises as the price of the good increases. Rather that each agent's willingness to pay for the good increases as other agents also consume the good (or express the desire to do so).

2 Evidence: Duflo and Saez (2003)

There is precious little credible work on social interactions. Duflo and Saez is an excellent example of a randomized experiment that is able to identify social interactions—though as they point out, the experiment is still under-identified in that only half of the behavioral parameters of interest can be estimated.

- Every employee can contribute to a Tax Deferred Account (TDA) up to $10,500.
- Benefits fair annually. Notification one month before.
- Participation rate is 34%, which is low relative to other universities.
- Encouragement design: offering a randomly chosen subset of employees a small amount of money for attending the fair.
- Randomization is at two levels: at level of department and at level of individuals in the department.
- Thus, two treatments: receiving a letter; being in a department where someone receives a letter. The treatment groups are: 11 equals receiving a letter and being in a department that someone received a letter (these have to go together); 10 not receiving a letter but being in a treated department; 00 no letter in department.
- Outcomes are: fair attendance, TDA participation after 4.5 months, TDA participation after 11 months.
- Table I makes it clear that *something* happened as a result of treatment. Fair attendance was 5 times as high among those who received the letter (group 11) as those who did not, and it was 3 times as high among those in departments where *someone else* received

a letter (group 11). Eleven months after the intervention, participation in the TDA was somewhat higher for both group 10 and group 11, though they do not appear different from one another.

- Table II presents reduced form estimates:

$$\begin{aligned} f_{ij} &= \alpha_1 + \beta_1 D_j + \varepsilon_{ij}, \\ Y_{ij} &= \alpha_2 + \beta_2 D_j + \omega_{ij}, \end{aligned}$$

where D is a dummy variable for being a member of a treated department.

- Notice the puzzle here. The D treatment clearly raises the probability of fair attendance and TDA enrollment. Those who received letters were more likely to attend the fair but less likely to enroll than coworkers in their departments who did not receive a letter.

2.1 Interpretation

There is no entirely unambiguous way to interpret theses results. Assumptions are needed.

- Consider the following equation for TDA participation:

$$y_{ij} = \alpha + \gamma_i f_{ij} + \Gamma D_j + u_{ij},$$

where D_j is a dummy indicating that the department was treated and f is an indicator equal to one if the i attended the fair. The fact that γ is indexed by i means that this is a random coefficients model.

- There are three treatment groups, corresponding to the combination of D and L received: $\{D, L\} \in \{00, 10, 11\}$. Note that there is no 01 group since if an individual received the letter, his department is also treated by definition.

- Define potential outcomes for fair attendance as $f(11), f(10), f(00)$

- Consider the following identification assumptions for fair attendance:

1. Monotonicity: For all i, $f_{ij}(11) \geq f_{ij}(10) \geq f_{ij}(0)$. So, departmental treatment weakly increases the probability of attending the fair, and receiving the letter weakly increase it further for each participant. Given monotonicity, we have the following potential outcome groups: [Is it reasonable to doubt monotonicity here? Does it matter if it's violated?]

   (a) Never takers: $f(11) = f(10) = f(00) = 0$.

   (b) Financial award compliers: $f(11) = 1 > f(10) = f(00) = 0$.

   (c) Social interaction compliers: $f(11) = f(10) = 1 > f(00) = 0$

   (d) Always takers: $f(11) = f(10) = f(00) = 1$.

2. Exclusion (for the second stage condition): u_{ij} independent of L_{ij} and D_j. This assumption must be interpreted with care. It says that (1) receipt of the letter does not *directly* affect TDA participation unless the participant attends the fair, and (2) that receipt of a letter by *someone* in the department does *not* affect TDA participation of non-recipients. The spillover, if it occurs, only works through the attendance of others in the department at the fair. Thus, a person in department j does not have to attend the fair to be affected by D_j. But someone in j needs to be induced to attend the fair by L_{ij} for D_j to be equal to 1.

- Thus, we have three parameters to estimate so far:

  $E[\gamma_i | f_{ij}(11) - f_{ij}(10) = 1]$ (average treatment effect for financial reward compliers)

  $E[\gamma_i | f_{ij}(10) - f_{ij}(00) = 1]$ (average treatment effect for social interaction compliers)

  and Γ, which is the 'social network effect parameter.'

- It's clear, however, that this setup is going to produce some puzzling results. We know that letter recipients were *less likely* to enroll in the TDA than non-recipients in their departments. Thus, the model will lead to the conclusion that the second parameter is less positive than the first. Citing the psychology literature, the authors note that receipt of a financial award may be de-motivating. For example, workers induced to go to the fair due to the promised $20 may be more likely *not* to take the information seriously.

- Assume that:

$$\gamma_i = \gamma_i^S - \nu L_{ij},$$

where γ_i^S is the standard treatment effect and ν is the de-motivating effect. This assumption is obviously motivated by the observation that those who received the letter were somewhat less likely to participate in the TDA than those who were in treated departments but did not receive a letter. Now we have four parameters to estimate (ATE for financial reward compliers, ATE for social interaction compliers, Γ, and ν), but only two treatments, L and D.

- Table III works through potential interpretations under different assumptions:

1. If we assume that γ is constant and that $\nu = 0$. we can identify γ and Γ with 2SLS by instrumenting f and D with the random assignments. (Note that we are still assuming the exclusion restriction above, which is not a trivial condition.)

2. Assume instead that Γ is zero and that $\nu = 0$, we can estimate $E\left[\gamma_i | f_{ij}(11) \geq f_{ij}(10)\right]$, the treatment effect for financial incentive compliers, by instrumenting f in the sample of treated departments. There is no TDA treatment effect for this group. Why? Because we are contrasting them with other workers in their departments whose behavior was affected by the spillovers. Thus, this contrast probably yields the 'wrong' answer.

3. We can estimate $E\left[\gamma_i | f_{ij}(10) \geq f_{ij}(00)\right]$, the treatment effect for social interaction compliers, by instrumenting for f using D for the subset of participants who did not receive a letter. However, it is probably not realistic to assume that the only 'treated' members of the department are those who attend the fair. Thus, this point estimate is likely an upper bound on the person-level causal effect.

4. We can potentially compare columns (2) and (3) of Table III to estimate ν if we are willing to assume that $E\left[\gamma_i | f_{ij}(11) \geq f_{ij}(10)\right] = E\left[\gamma_i | f_{ij}(10) \geq f_{ij}(00)\right]$.

5. Finally, the 'naive IV' that does not account for social interactions in treated departments would imply that the experiment did not raise TDA participation. This is because many of the controls are indirectly treated.

Conclusions: (1) It's hard to do this type of exercise well; (2) This is a successful effort.

3 Montgomery, 1991: Social Networks and Labor Market Outcomes

Most people find jobs through social networks. Many employers believe that employee referrals are a useful device for screening job applicants. The 1991 Montgomery paper offers an explanation for why well-connected workers may fare better than poorly connected workers, and why firms hiring through referral might earn higher profits. This is widely considered a seminal paper. I am not aware of empirical research that directly builds from this model, however.

3.1 Workers

- There are two periods.
- Each worker lives for one period.
- There are two types of workers, both equally populous: H and L.
- High ability workers produce 1 and low ability workers produce 0.
- Employers cannot observe worker ability prior to hiring.
- There are no output-contingent contracts. (As Montgomery notes, if such contracts were readily feasible, the entire issue of screening and the considerable effort that employers spend on worker selection would seemingly be pointless.)

3.2 Firms

- Each firm may employ up to 1 worker.
- Profit equals productivity minus the wage.
- Product price of 1 is exogenous determined.
- Firms must set wages prior to learning the productivity of their workers.
- They do observe their own worker's type prior to the start of the second period (of course, that worker is about to die).

- There is free entry

3.3 Social structure

- Each period 1 worker knows (is 'tied' to) at most one period 2 worker. Specifically, the probability of such a tie is $\tau \in [0, 1]$. So, if there are $2N$ workers in period 1, there will be $2\tau N$ ties.

- Ties are drawn with replacement. Thus, although a period 1 worker can only have one tie, a period 2 worker can have any non-negative number of ties. *Hence, the distribution of ties in period 2 is binomial with mean τ and variance $\tau(1-\tau)$.*

3.4 Two observations

1. Hiring a worker in period 1 has an associated option value. This arises from the fact that the period 1 worker may be type H *and* may have a tie to another type H worker (the joint probability being $\frac{1}{2}\alpha\tau$).

2. Intuition should suggest that this network mechanism could give rise to adverse selection.

    - An employer who observes that she has hired a type H worker in period 1 who has a network tie (joint probability $\frac{1}{2}\tau$) faces probability $\alpha > 1/2$ of receiving a referral to another type H worker.

    - This conditional probability of α is better than the employer could do by chance *in period 1* (the unconditional probability being 1/2).

    - This conditional probability is *better still* than the employer could obtain by chance in period 2. This is because, as will be shown, type H period 2 workers will be more likely than type L workers to receive offers via referrals (and to accept those offers).

    - Consequently, the pool of workers on the open market (i.e., those not hired through referrals) will be adversely selected. The workers available on the period 2 open market will have $\Pr(H) < \frac{1}{2} < \alpha$.

3.5 Main proposition

The solution to the model is non-trivial and enlightening. Montgomery starts with the following proposition, which is initially assumed and subsequently proved:

Proposition 1 *A firm makes a referral offer iff it employs a type H worker in period 1. Referral wage offers are dispersed over the interval $[w_{M2}, \bar{w}_R]$.*

In this proposition, w_{M2} is the market wage in period 2.

3.6 Structure of the proof

The structure of the proof is complex and features many twists and turns. Here's the roadmap:

1. We first solve the worker's problem
   (a) Calculate the structure of ties (i.e., how many ties a worker can expect to have)
   (b) Calculate the probability that an offer is accepted by a period two H worker
2. Calculate the market wage for workers who receive no referral.
3. Given the market wage and the expected productivity of referrals, calculate firms' referral wage offers (i.e., the offers that they make if they hire an H worker in period 1 and that worker has a network tie). Given the proposition above:
   (a) Show that offering W_{M2} is one feasible strategy
   (b) Calculate $\bar{W}_R$ and show that offering a higher wage than $\bar{W}_R$ is not an equilibrium strategy
   (c) Argue that wage offers on the interval $\left[W_{M2}, \bar{W}_R\right]$ must all be equally good
4. Prove that firms who have hired L workers hire on the open market (i.e., not through referrals)
5. Calculate *period one* wages by backward induction
6. Reflect on our accomplishment

3.7 The worker's problem

3.7.1 The structure of ties

- The distribution of ties in period 2 is binomial with mean τ and variance $\tau(1-\tau)$. The way to think about the problem is that there are a total of $2N$ draws with replacement from the urn of period 2 workers, where each draw selects one period 2 worker with probability τ.

- The likelihood that a given period 2 worker receives exactly k ties is:

$$\Pr[k_i = k] = \binom{2N}{k}\left(\frac{\tau}{2N}\right)^k\left(1-\frac{\tau}{2N}\right)^{2N-k},$$

  where

$$\binom{2N}{k} = \frac{2N!}{k!\,(2N-k)!}.$$

- Now assume that ties exhibit *homophily.* Conditional on a tie existing, the probability that the period 2 worker is of the same type as the period 1 worker to which he is tied is $\alpha > 1/2$.

- Thus, if a firm observes that its period 1 worker is type H (which occurs in N cases), there is a probability τ that this worker has a tie, and a probability $\alpha\tau$ that the worker has a tie to a period 2 worker of type H.

- The probability that a given **period** 2 worker of type H has k ties to period 1 workers of type H is:

$$\Pr[k_i = k] = \frac{(N-k)!}{k!\,(N-k)!}\left(\frac{\alpha\tau}{N}\right)^k\left(1-\frac{\alpha\tau}{N}\right)^{N-k}.$$

  Note that the $2N$ term has become an N term because there are only N type H workers in the total population of $2N$ workers.

3.7.2 The probability that a period two H worker accepts a wage offer

- Consider a given type H worker in period 2. Assuming the proposition holds, so all referral offers exceed the market wage. The probability that H accepts a given referral wage offer

w_{Ri} from firm i is the probability that the worker receives no better offer from some other firm $\dot{j}$:

$$\begin{aligned} \Pr\left[H \text{ accept } w_{Ri}\right] &= \Pr[H \text{ receives no offer higher than } w_{Ri}\ \forall \text{ firm } j \neq i] \\ &= \Pi_{j\neq i}\left\{1-\Pr\left[H \text{ receives offer } w_{Rj} > w_{Ri}\right]\right\}. \end{aligned}$$

- The probability that a *given* firm j makes a better wage offer to H is:

$$\begin{aligned} \Pr\left\{H \text{ receives an offer } w_{Rj} > w_{Ri}\right\} &= \Pr\left\{\text{firm } j \text{ makes an offer to } H\right\} \\ &\quad \times \Pr\left\{w_{Rj} > w_{Ri}\right\}. \end{aligned}$$

- Given that there are $2N$ workers in period 1, it must be that N firms employ type H workers. Assuming (per the proposition) that each firm chooses its referral wage offer by randomizing over the equilibrium wage distribution $F(\cdot)$, then

$$\Pr\left\{H \text{ receives an offer } w_{Rj} > w_{Ri}\right\} = \frac{\alpha\tau}{N}\left[1-F\left(w_{Ri}\right)\right],$$

for all firms j employing a high-ability worker in period 1, where $F(\cdot)$ is the cumulative density of wage offers, and $\alpha\tau/N$ is the probability that H has a tie to a worker at firm j.

- The probability that w_{Ri} is the maximum of all wage offers received by this H worker is therefore:

$$\Pr\left\{H \text{ accepts } w_{Ri}\right\} = \left(1-\left(\frac{\alpha\tau}{N}\right)\left[1-F\left(w_{Ri}\right)\right]\right)^{N-1}.$$

- To make further progress on this expression, we need to use the Bernoulli formula. What we want to calculate is the probability that no offer received by worker H exceeds w_{Ri}. Again, imagine that there are N draws from the urn of period 2 workers (not $2N$ because the relevant set draws from type H period 1 workers). The probability that any specific draw results in a wage offer to a given H worker exceeding w_{Ri} is $\frac{\alpha\tau}{N}\left[1-F\left(w_{Ri}\right)\right]$ .Thus, the process of receiving superior wage offers is also Bernoulli. (This is an important subtlety: it's not just the frequency of ties that is Bernoulli but also the frequency of superior wage offers.)

- The probability that a period two H worker receives exactly k superior wage offers given w_{Ri} is (we use $N-1$ because one offer is already in hand):

$$\Pr\left[k_i = k | w_{Ri}\right] = \frac{(N-1-k)!}{k!\,(N-1-k)!}\left(\frac{\alpha\tau\left[1-F\left(w_{Ri}\right)\right]}{N}\right)^k\left(1-\frac{\alpha\tau\left[1-F\left(w_{Ri}\right)\right]}{N}\right)^{N-1-k}.$$

- We need to calculate the probability that an H worker receives zero superior wage offers. Why? Because a worker who receives zero superior offers will take the offer that is in hand

$$\Pr\left[k_i = 0 | w_{Ri}\right] = \left(1-\frac{\alpha\tau\left[1-F\left(w_{Ri}\right)\right]}{N}\right)^{N-1}.$$

3.7.3 The tricky part

- This expression is very close to the following limit:

$$\lim_{n\to\infty}\left(1-\frac{\alpha\tau\left[1-F\left(w_{Ri}\right)\right]}{n}\right)^n = \frac{1}{e^{\alpha\tau[1-F(w_{Ri})]}} = e^{-\alpha\tau[1-F(w_{Ri})]}.$$

- More familiar:

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e},$$

and

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e.$$

- So, putting pieces together. The probability that H accepts firm $i's$ offer is:

$$\Pr\left[H \text{ accepts } w_{Ri}\right] = e^{-\alpha\tau[1-F(w_{Ri})]},$$

and similarly, the probability that L accepts firm $i's$ offer (remember, period 2 workers are also tied to period 1 H workers with probability $(1-\alpha)\,\tau$) is:

$$\Pr\left[L \text{ accepts } w_{Ri}\right] = e^{-(1-\alpha)\tau[1-F(w_{Ri})]}.$$

3.8 The firm's problem

3.8.1 Period 2 *market* (non-referral) wages

- What should a firm offer to a period two worker who is in the open market (i.e., not hired by through a referral)?

- Since referral wage offers always dominate market wage offers, the probability that a type H worker enters the market in period 2 is equal to the probability of receiving no wage offers:

$$\Pr\{\text{Market } |H\} = e^{-\alpha\tau}.$$

and for type L workers, this probability is:

$$\Pr\{\text{Market}|L\} = e^{-(1-\alpha)\tau}.$$

- This brings us to the adverse selection problem. The offer wage (equal to expected productivity) of period 2 workers on the market is:

$$w_{M2} = \Pr\{H|\text{Market}\} = \frac{e^{-\alpha\tau}}{e^{-\alpha\tau} + e^{-(1-\alpha)\tau}} < \frac{1}{2}.$$

Period 2 market workers are therefore adversely selected.

- Moreover:

$$\frac{\partial w_{M2}}{\partial \tau} < 0, \frac{\partial w_{M2}}{\partial \alpha} < 0.$$

The adverse selection problem becomes more severe the denser are social networks ($\tau+$) and the greater is the degree of homophily ($\alpha+$).

3.8.2 Period 2 *referral* wages

- Consider the expected period 2 profit of a firm employing a type H worker in period 1 and setting a referral wage of w_R.

$$\begin{aligned} E\Pi_H(w_R) &= \Pr[\text{Type } H \text{ referral hired}|w_R] \cdot (1 - w_R) \\ &\quad + \Pr[\text{Type } L \text{ referral hired}|w_R] \cdot (-w_R) \\ &\quad + \Pr[\text{No referral hired}|w_R] \cdot 0, \end{aligned}$$

where in this expression we use the fact that hiring in the outside market must have expected profit of zero (i.e., it's a competitive situation—unlike referrals where markets are 'thin').

- Conditional on hiring a type H worker in period 1, the probability of hiring a period two H worker with offer wage w_R is:

$$\Pr\left[\text{Type } H \text{ referral hired}|w_R\right] = \alpha\tau e^{-\alpha\tau[1-F(w_R)]}.$$

Similarly,

$$\Pr\left[\text{Type } L \text{ referral hired}|w_R\right] = \left(1-\alpha\right)\tau e^{-(1-\alpha)\tau[1-F(w_R)]}.$$

So,

$$E\Pi_H(w_R) = \alpha\tau e^{-\alpha\tau[1-F(w_R)]} \times (1-w_R) + \left(1-\alpha\right)\tau e^{-(1-\alpha)\tau[1-F(w_R)]} \times \left(-w_R\right).$$

- In equilibrium, this profit level must be robust to deviations.

3.8.3 The lower-bound wage

- A firm could simply offer the market wage w_{M2} (or $w_{M2}+\varepsilon$) to a referral (a 'low-ball' strategy), and that referral will be accepted if the worker has received no other offers—which will occur with positive probability, even in an infinitely large workforce, given that there are fewer than half as many referral offers as workers.

- Given that H worker is receiving a referral offer, what is the probability that he receives no other?

$$\begin{aligned}
\Pr\left[k_i = 1|H, k_i \geq 1\right] &= \frac{(N-1-k)!}{k!\,(N-1-k)!}\left(\frac{\alpha\tau}{N}\right)^k\left(1-\frac{\alpha\tau}{N}\right)^{N-k-1} \\
&= \left(1-\frac{\alpha\tau}{N}\right)^{N-1} \\
&\approx e^{-\alpha\tau}
\end{aligned}$$

- Similarly, for an L worker:

$$\Pr k_i = 1|L, k_i \geq 1 \approx e^{-(1-\alpha)\tau}$$

- So, the profitability of the low-ball strategy is:

$$\begin{aligned}
E\Pi_H(w_{M2}) &= \alpha\tau e^{-\alpha\tau}(1-w_{M2}) + (1-\alpha)\tau e^{-(1-\alpha)\tau}(-w_{M2}) \\
&= \alpha\tau e^{-\alpha\tau}\left(1-\frac{e^{-\alpha\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}}\right) + (1-\alpha)\tau e^{-(1-\alpha)\tau}\left(-\frac{e^{-\alpha\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}}\right) \\
&= \left(\alpha\tau e^{-\alpha\tau} - \frac{\alpha\tau e^{-2\alpha\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}}\right) + (\alpha-1)\left(\frac{\tau e^{-(1-\alpha)\tau-\alpha\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}}\right) \\
&= \frac{\alpha\tau e^{-\alpha\tau}\left(e^{-\alpha\tau}+e^{-(1-\alpha)\tau}\right) - \alpha\tau e^{-2\alpha\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}} + \frac{(\alpha-1)\tau e^{-\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}} \\
&= \frac{\alpha\tau e^{-\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}} + \frac{(\alpha-1)\tau e^{-\tau}}{e^{-\alpha\tau}+e^{-(1-\alpha)\tau}} \\
&= \frac{(2\alpha-1)\tau}{e^{-\alpha\tau+\tau}+e^{-(1-\alpha)\tau+\tau}} \\
&= \frac{(2\alpha-1)\tau}{e^{\alpha\tau}+e^{(1-\alpha)\tau}}.
\end{aligned}$$

- Call this profit level c:

$$c(\alpha,\tau) \equiv \frac{(2\alpha-1)\tau}{e^{\alpha\tau}+e^{(1-\alpha)\tau}} > 0,$$

which is positive since $\alpha > 1/2$. By implication, firms who hire a type H worker in period 1 earn a positive expected profit in period 2. Moreover,

$$c_\alpha > 0, c_\tau > 0.$$

3.8.4 The upper-bound wage

- Clearly, in equilibrium, firms with a type H worker must receive $c(\alpha,\tau)$ from either making an offer of w_R or making an offer of w_{M2}. So, it must be the case that for all w_R :

$$\begin{aligned}
c(\alpha,\tau) &= \alpha\tau e^{-\alpha\tau[1-F(w_R)]}\times(1-w_R) \\
&\quad + (1-\alpha)\tau e^{-(1-\alpha)\tau[1-F(w_R)]}\times(-w_R) \\
\forall w_R &\in [w_{M2},\bar{w}_R]
\end{aligned}$$

Montgomery offers the interpretation that either firms randomize their offers over the entire distribution or else a fraction $f(w_R)$ of firms offers each wage for sure. Montgomery claims (and I, for one, believe him) that this expression does not have a closed form solution.

- However, we know that if a firm offers $\bar{w}_R$, it hires the referred worker for sure (conditional on having a H worker with a tie). Thus, we can solve for $\bar{w}_R$

$$\begin{aligned}
\frac{(2\alpha-1)\tau}{e^{\alpha\tau}+e^{(1-\alpha)\tau}} &= \alpha\tau\left(1-\bar{w}_R\right)+(1-\alpha)\tau\left(-\bar{w}_R\right) \\
\frac{(2\alpha-1)\tau}{e^{\alpha\tau}+e^{(1-\alpha)\tau}} &= \alpha-\alpha\bar{w}_R-\bar{w}_R+\alpha\bar{w}_R \\
\bar{w}_R &= \alpha-(2\alpha-1)/\left(e^{\alpha\tau}+e^{(1-\alpha)\tau}\right) \\
\bar{w}_R(\alpha,\tau) &= \alpha-c(\alpha,\tau)/\tau.
\end{aligned}$$

- One can easily show that $\partial\bar{w}_R/\partial\alpha > 0$, $\partial\bar{w}_R/\partial\tau > 0$. That is, the maximum referral wage offer is also increasing in network density and homophily.

- Summing up, the expected period 2 profit from hiring a type H worker is:

$$E\Pi_H = (\alpha-\bar{w}_R)\tau,$$

  which is also increasing in α and τ.

3.9 Do firms who have hired L workers hire *only* on the open market?

- We have established that firms holding type H workers who have ties make offers on the interval $[w_{M2},\bar{w}_R]$. Lower offers are not accepted and higher offers do not increase the probability of attracting a worker.

- We now need to prove that firms hiring type L workers in period 1 will hire through the market.

- Imagine a firm deviated from the equilibrium by making a referral offer $w_{M2} < w_R < \bar{w}_R$ via its type L worker:

$$E\Pi_L(w_R) = (1-\alpha)\tau e^{-\alpha\tau[1-F(w_R)]}\times(1-w_R)+\alpha\tau e^{-(1-\alpha)\tau[1-F(w_R)]}\times(-w_R).$$

- It is immediately apparent that

$$\frac{E\Pi_L(w_R)}{\partial w_R} < \frac{E\Pi_H(w_R)}{\partial w_R}.$$

This follows because any action that increases the chance of the offer being accepted increases the winner's curse (since odds are better than 50% that the worker will be type L). And we know by construction that

$$\partial E\Pi_H (w_R) / \partial w_R = 0 \; \forall \; w_R \in [w_{M2}, \bar{w}_R] .$$

Thus,

$$\frac{\partial E\Pi_L (w_R)}{\partial w_R} < 0$$

.

- By implication, $E\Pi_L (w_R)$ is maximized at $w_R = w_{M2}$. But we can show that this expectation is negative using the formula derived above for w_{M2}.

$$\begin{aligned} E\Pi_L (w_{M2}) &= (1-\alpha) \tau e^{-\alpha\tau[1-F(w_R)]} \times (1 - w_{M2}) + \alpha\tau e^{-(1-\alpha)\tau[1-F(w_R)]} \times (-w_{M2}) \\ &= (1-\alpha) \tau e^{-\alpha\tau[1-F(w_R)]} \times \left(1 - \frac{e^{-\alpha\tau}}{e^{-\alpha\tau} + e^{-(1-\alpha)\tau}}\right) \\ &\quad + \alpha\tau e^{-(1-\alpha)\tau[1-F(w_R)]} \times \left(-\frac{e^{-\alpha\tau}}{e^{-\alpha\tau} + e^{-(1-\alpha)\tau}}\right) \\ &= \frac{(1-2\alpha) e^{-\tau}}{e^{-\alpha\tau} + e^{-(1-\alpha)\tau}}. \end{aligned}$$

This expression is negative since $\alpha > 1/2$. Thus, a firm employing a L worker in period 1 will offer the market wage in period 2. It will *not* use referrals.

3.10 Calculating period 1 wage

- Finally, consider the period 1 wage (where types are not known). This wage incorporates both the expectation of period 1 productivity and the option value of a period 2 referral:

$$w_{M1} (\alpha, \tau) = \frac{1}{2} + \frac{1}{2} c (\alpha, \tau) = \frac{1}{2} [1 + c (\alpha, \tau)] .$$

Given prior results on $c(\cdot)$, this implies that w_{M1} is increasing in α and τ. That's an interesting result in that low-ability workers benefit from the uncertainty in period 1 but are harmed by the adverse selection in period 2.

3.11 Main results: Summary (AKA, reflect on our accomplishment)

The key results of the model are:

1. In equilibrium, each worker's wage is determined *not* by his actual skill but by the number and types of social ties he holds (though of course these are not independent of worker skill). Period 2 workers with more ties to high-ability Period 1 workers receive more referrals and thus higher expected wages. Workers with no ties to high ability workers find employment on the open market, which is afflicted with adverse selection. (Thus, *It's who you know, not what you do.*)

2. In equilibrium, workers obtained through referrals are of higher quality. So, it is rational for firms to prefer referrals, and these referrals to generate rents (which are dissipated into the labor market in period 1 in the form of option value payments).

3. Both α and τ have similar effects in the model (and it's not really clear that these parameters need to be conceptually distinct). An increase in either raises the top wage ($\bar{w}_R$) and lowers the bottom wage (w_{M2}) and also raises the period 1 wage (w_{M1}).

4. One model extension in which α and τ might differ is in a setting where there are different groups that are intrinsically of equal ability and have similar homophily along the lines of ability but one group has a better network ($\tau+$) than the other. So, take the case of Blacks and Whites. *If Whites have a better network, even with equal ability, the option value of hiring Whites is greater than that of hiring Blacks in period 1, even if the White network advantage stems from their connection to type H Blacks. Thus, network density redistributes income from the those who are referred to those doing the referring.* Of course, referred workers also benefit. The more offers a period 2 worker receives, the higher his wage. (Remember that the accepted wage is the highest wage offered, and each offer is a random variable. So a larger number of offers increases the expected accepted wage).

5. If there is a complementarity between ability and the production technology (such that firms would use one technology if they could be reasonably confident of getting type H

workers and another if they could not), then an increase in τ might improve productive efficiency but also increase inequality. This idea is the essence of the model in Acemoglu 1999 in the *AER*.

This is a path-breaking paper that continues to be widely cited. What is most thought-provoking about it is: (a) it suggests a non-market mechanism that could affect economic outcomes in a competitive environment (assumption: there is *not* a market for social ties); (b) it provides the insight that the option value of networks accrues (at least in part) to those at the top of the network (referring) rather than exclusively those being referred.

4 The effect of Social Networks on Employment and Inequality: Calvo-Armengol and Jackson (2004)

The 2004 paper by Calvo-Armengol and Jackson presents a distinctly different way of modeling social networks and their potential impact on employment. In their model, there are no prices, no wages, no adverse selection and no individual heterogeneity—so the market mechanism is quite impoverished relative to Montgomery (1991). However, the network structure is richer. In particular, it has a topological aspect wherein which workers are connected to one another via multiple nodes. The model appears to capture the seemingly important phenomenon that individuals are tied to one another both directly (as in τ in Montgomery) and *indirectly* through other acquaintances. This extremely sparse model gives rise to surprisingly rich interactions.

4.1 Model basics

- There are n agents.
- Time involves in discrete periods indexed by t
- The vector $\mathbf{s}_t$ describes the employment status of all agents at time t.
- If agent i is employed at the end of period t, then $s_{it} = 1$, and if unemployed, $s_{it} = 0$.
- A period t begins with some agents employed and others not, described by the status $\mathbf{s}_{t-1}$ from the last period.

- Next, information about jobs opening arrives.
- In each period, any given agent hears about a new job opening with probability $a \in (0,1)$.
- If the agent is unemployed, she takes the job.
- If the agent is employed, she passes the information to someone in her immediate network who is unemployed.

4.2 Network structure

- Any two agents either know one another or not.
- Information only flows between agents who know each other.
- A graph g summarizes the links of all agents, where $g_{ij} = 1$ indicates that i and j have a link.
- Links are reflexive: $g_{ij} = g_{ji}$.
- If an agent hears about a job and is employed, she passes the information to another randomly chosen *linked* agent who is unemployed.
- If all of her linked agents are employed, the information is lost.
- Thus, the probability that agent i learns about a job *and* that this job is taken by agent j is described by $p_{is}(\mathbf{s})$:

$$p_{is}(\mathbf{s}) = \begin{cases} a & \text{if } s_i = 0 \text{ and } i = j \\ a\left(\sum_{k:s_k=0} g_{ik}\right)^{-1} & \text{if } s_i = 1,\ s_j = 0, \text{ and } g_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} .$$

- At the end of a period, each employed worker loses her job with probability b.

Some notes:

1. If there were no network structure, this setup be a simple Markov process with transition probabilities $\Pr(U \to E) = a$, $\Pr(E \to U) = b$, and steady state employment probability of $a/(a+b)$. [In steady state $E \times b = U \times a$. So, $\frac{E}{U} = \frac{a}{b}$, and the employment to population rate is $E/(E+U) = a/(a+b)$.

2. There is no communication beyond one node in the graph. That is, a job referral does not pass from agent i to j to k.This means that any network effects beyond immediate acquaintances are indirect.

3. If we think about two agents each linked to a third agent, these two agents are competitors for job information in the short run. But in the longer run, they help to keep one another employed indirectly. In particular, if the shared friend becomes unemployed, then each acquaintance potentially helps to re-employ the shared friend. This benefits the other acquaintance because if the shared friend is re-employed, he potentially helps the acquaintances should they become unemployed.

4.3 Some basics

The paper is chock-full of thought-provoking examples.

- Figure 2 shows that although agents are competitors for the job info possessed by a mutual friend, there will still tend to be a positive correlation in employment rates between agents with a mutual friend.

- A formal proposition shows that under fine enough time subdivisions (where a and b are divided by some common T, so the periods become extremely short), the unique steady-state long-run distribution on employment is such that the employment statuses of *any* path-connected agents are positively correlated (where path-connectedness can mean a direct or indirect connection).

- Moreover, the positive correlation holds across any arbitrary time span. That is, agent $i's$ employment status at time t is correlated with agent $j's$ status at time t' for all t and t'.

4.4 Duration dependence

- Suppose that $a = 0.1$ and $b = 0.015$. Given that a person has been unemployed for at least each of the last X periods. What is the probability that she will find employment this period?

- Figure 6 illustrates. Notice the negative duration dependence of re-employment. The longer an agent is unemployed, the lower his re-employment probability. Yet we know there is no heterogeneity among agents, and there is no change in aggregate labor market conditions. What's going on?

- The longer that i has been unemployed, the higher the expectation that $i's$ connections and path connections are themselves unemployed. This makes it more likely that $i's$ connections will take jobs that arrive rather than passing them down.

- To clarify, this is *not* a causal effect of i being unemployed. Rather, i having remained unemployed for some time provides information about the poor employment status of his connections. Thus, our expectation of his exit hazard falls.

- This observation interesting because it suggests that individuals may experience time-varying serially correlated reemployment probabilities that are explained by network status but not otherwise due to either person-level characteristics or the aggregate state of the economy. So, if my friends lose a job due a plant closing, I may be less likely to find re-employment conditional on job loss not because fewer jobs are available but because I'm less likely to hear about them.

4.5 Dynamics

- Given the network externalities in the model, it should be clear that the state of aggregate employment can cycle. If the network gets close to full employment, unemployed agents become ever more likely to find jobs. If employment falls due to a set of chance events, each newly unemployed worker also becomes less likely to find reemployment. This leads to booms and busts.

- Nevertheless, the sharing of information about jobs means that more offers get used (fewer get lost) than in a network with no connections. So, aggregate employment is likely to remain higher in a networked than non-networked market. See Figure 7.

4.6 Dropping out and contagion

Perhaps the most interesting part of the article in my mind is the section on labor market dropouts. As constructed so far, the steady states of a network does not depend on initial conditions. That is, whether all nodes are initially employed or unemployed, the model will move towards the same average level of employment. If the model is expanded to allow labor market drop-out, however, the dynamics get richer.

In particular, CJ add a 'drop out' option as an absorbing state of the model. Agents face a present value of costs $c_i \geq 0$ of remaining in the market. The outside option is zero. So an agent will only remain in the LF if the discounted expected value of wages exceeds c_i. The per-period wage will be fixed at 1. A key assumption is that once an agent drops out, he no longer passes along job referrals to linked agents and yet remains in the network. Thus, his a signals are effectively lost. Moreover, it is assumed for simplicity that the dropout's connected agents still send referrals to him (and thus they are squandered). [Without this assumption, the network structure would effectively change when an agent dropped out, which would complicate things greatly.]

Clearly, dropout probabilities will be rising in costs and declining in wages. More interestingly, an agent's dropout probability will depend on her network. The better a person's network, the greater the re-employment hazard following job loss, and thus the greater the discounted expected future value of earnings.

Most interestingly, agents dropout decisions can have *contagion* effects because one agent's exit weakens the network for each of her path-dependent agents. Accordingly, decisions to remain in the labor force are *strategic complements*—the more participants remain in, the more advantageous it is for a given agent to remain. The dropout game is therefore *supermodular* (in calculus terms, positive cross-partials for all components). This means the game has a *maximal equilibrium* in pure strategies where the set of agents staying in the game is maximized (though the specific identities of the agents could change across equilibria or points in time). Dropouts can have negative contagion effects due to the strategic complementarities among agents' actions.

To simulate this setting, CJ set the cost c_i uniformly on $[0.8, 1]$ and fix the wage at 1. They assume a complete network—every agent is linked. The discount factor is 0.9 and the transition probabilities are $a = 0.1$ and $b = 0.015$.

To calculate contagion, they first calculate drop out rates that would occur if every agent assumed that every other agent were staying in the market. Then they calculate how many agents drop out in equilibrium (calculated by simulation). The difference between these two is the estimated contagion effect.

See Tables 2 and 3. Several things are interesting about these results:

- Contagion effects are substantial in some cases.
- With a very large number of nodes in a fully connected network, the contagion effect becomes negligible. This is because any one person's dropout decision is inconsequential for all other agents.
- Contagion effects are more important when workers start unemployed than employed—presumably because more workers drop out immediately, leading to further dropout. Thus, the network is state dependent in that the starting state influences the long-run equilibrium. A proposition demonstrates this observation formally. If the starting state in two identical networks is person-by-person (weakly) higher in one network than the other, then the network that starts initially higher has higher steady-state employment, and this is true for all agents in any component of the network for which equilibrium drop-out decisions differ across the two groups.
- A final observation is that holding the equilibrium unemployment rate constant $U = a/(a+b)$, a higher rate of turnover (higher b) induces a greater level of dropout. The reason is that the likelihood that an agent and his linked agents are simultaneously unemployed is higher in the setting where job destruction is more frequent. Thus, they are more likely to be in competition for job arrivals, which leads to an increase in dropout.

4.7 Substantive conclusions from CJ

I personally find it difficult to know how much stock to place in this thought-provoking paper. The stylized takeaway for me is that there could be important externalities in job awareness, leading to a dropout behavior that is contagious *not* due to peer imitation but simply due to adverse information externalities. The concern for me is that the model is so stripped-down that it's hard to call it an economic model:

- There are no wages (or they are parametric).
- Labor demand is independent of supply—so, if fewer agents work, this does not increase the arrival rate of offers or reduce the job destruction rate of remaining workers.
- Network structure is exogenous and fixed.
- Agents are not maximizing anything in particular—except when calculating drop-out.

Thus, this model is very much in the style of a Sante Fe Institute computerized-automata model: agents with a very limited repertoire of behavior are set loose in a simulated economic environment and then we study the emergent properties of that environment. When you read the Becker-Murphy (2000) volume on *Social Economics*, you will see that Becker-Murphy take issue with this type of modeling (though obviously not with this paper). This type of model will have to evolve considerably to make contact with more general economic models.