

Notes on Hedström and Åberg

Cosma Rohilla Shalizi

30 April 2008

Abstract

Notes on “Quantitative research, agent-based modelling and theories of the social”, by Hedström and Åberg, chapter 6 (pp. 114–144) in Hedström (2005). For 36-462.

This chapter has both a theoretical/methodological aim, and an empirical one. The theoretical aim is to explain a non-trivial social phenomenon in what the authors regard as a methodologically sound way, to exemplify the virtue of the methodology. (In the book, this is what Hedström has spent the previous five chapters building up to.) The empirical aim is to study peer effects in youth unemployment.

Let’s take the theory side first. The picture they have in mind is something like this:

$$\begin{array}{ccc} \text{Macro}(t) & \xrightarrow{\text{prediction}} & \text{Macro}(t+1) \\ \uparrow \text{aggregation} & & \uparrow \text{aggregation} \\ \text{micro}(t) & \xrightarrow{\text{causation}} & \text{micro}(t+1) \end{array}$$

The desire is to fill in the outer loop, rather than relying on phenomenological models of the upper arrow. The latter are basically unstable, since they rely on many small-scale details not changing very much. There will of course be exceptions, where the macroscopic variables are very specifically defined, but in general arbitrarily-chosen macro variables won’t even form a Markov process, let alone a causal Markov process (Shalizi and Moore, 2003).

Micro level processes should (say Hedstrom and Aberg) reflect the *choices* of the agents among their available options, in turn coming from beliefs and desires, and how the consequences of those choices create conditions for others’ choices.¹

Let’s turn specifically to unemployment. How can we think about its dynamics while keeping our eye on Hedström’s trinity of “desires, beliefs, opportunities”? You need to look in order to find a job; so looking has to be better than not looking. Also, if you do look, you need information to actually find the job. The authors draw a diagram from this; unfortunately they get it wrong! Here is what their figure 6.2 should look like.

¹That last is the only sensible meaning of “dialectic” (Kautsky, 1988; Elster, 1982, 1985).

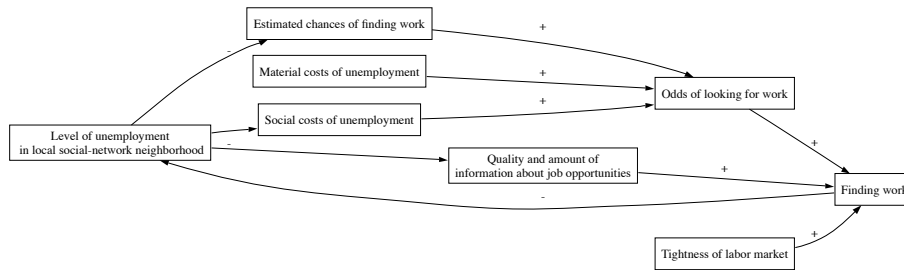


Figure 1: Feedback diagram replacing figure 6.2 in the chapter.

The reasoning is that having lots of the people you know lowers the social costs of being unemployed yourself (vernacularly: you feel like less of a slacker/loser), lowers your estimate of the odds of finding work (we mentally survey our environment), and reduces the quality and amount of the information you have about job openings. The first two channels act on whether you *choose* to look for work. If you do so choose, you still have to find it, and that's where information comes in. The feedback loop is closed by the fact that if you do find a job, you reduce your social neighbors' local unemployment levels.

Individuals' characteristics and interactions (finding or losing work, looking for jobs or not, etc.) thus aggregate up to the macro-level fluctuations in employment.

Data They had demographic information on all young adults age 20–24 in greater Stockholm, plus employment history, including week-by-week employment status, neighborhood of residence at end of each year, for years 1993–1999.

Model building Why not just directly implement the model in the figure? Because we don't have the measurements.

1. No measurement of the three variables in the channel from “local network unemployment” to “looking for work” — can't measure social costs, can't measure subjective odds, and can't measure quality of information. Also, we can't measure whether they looked for work! However, we can measure whether they found work. And all three channels have a negative impact on finding work. Solution: lump the three channels into a single channel with a negative sign, and hope this doesn't introduce too much systematic distortion.
2. No measurement of social networks. Solution: replace unemployment among network neighbors with unemployment among *geographic* neighbors. Hope that the two are strongly correlated. Strictly speaking this is an error-in-variables model now (as in Figure 2), which *as a rough rule of thumb* means that the actual impact of the predictor on the response is understated.

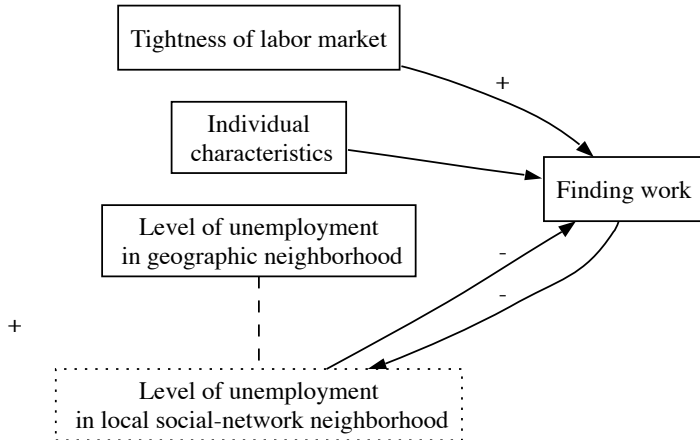


Figure 2: Reduced form of the model. Note that unemployment in a geographic neighborhood, which is observable, is used as a proxy for unemployment in the social-network neighborhood, which is unobservable (with this data). This distinction is elided when Hedström and Åberg actually do their estimation.

They use a logistic model for whether the unemployed found jobs or not. That is, they predict that the probability of unemployed individual i finding work at time-period t is

$$\frac{e^{\alpha + \beta U_{i,t} + \gamma \cdot \mathbf{Z}_{i,t}}}{1 + e^{\alpha + \beta U_{i,t} + \gamma \cdot \mathbf{Z}_{i,t}}} \quad (1)$$

where $U_{i,t}$ is the unemployment level at time t in i 's neighborhood, and $\mathbf{Z}_{i,t}$ is a vector of covariates for i at time t . The β term measures the strength of social influence. (This is not their notation.) Said differently, the probability is

$$\iota^{-1}(\alpha + \beta U_{i,t} + \gamma \cdot \mathbf{Z}_{i,t}) \quad (2)$$

where $\iota(p) = \log p / (1 - p)$ is the **logistic transform**. Written in this way, we can see that this is a **generalized linear model** (GLM), with the logistic transform as the **link function**. Notice that they provide no justification for either the logistic link function or for the linearity of the predictor.

They actually fit two variants of this model: one with just social influence (equivalently, γ fixed at 0), and one with a slew of demographic covariates added in. The latter fit much better, even allowing for the much larger number of parameters.

The estimated β is much smaller when covariates are added. This suggests *prima facie* a tendency for demographically-similar individuals to be found in the same districts. How much more of the apparent residual influence is due to similarity of *un*-observed but employment-relevant characteristics? (And how much of that similarity is due to processes of social influence?)

Phenomenology of the model They simulate the model on a homogeneous lattice, leaving out the demographic covariates. This produces clusters of persistent unemployment, even when not primed with such clusters. Since this is an Ising model below the critical temperature, that’s not surprising behavior. (They do not mention, and may not realize, that this is an Ising model.) At this level, the model is very similar to the model of persistent inequality in Oomes (2003).

They also simulate the model with the original geography and demographics, and find that you get bigger fluctuations as β increases, as one would expect — certainly bigger than if $\beta = 0$.

Checking the model The most basic check is the correlation between predicted and real unemployment levels: this is pretty good at around 0.8.

They also take selected *predicted* by-district unemployment levels, and look at the distribution of *actual* unemployment levels for those districts. These look OK by eye but no formal testing is done. We can think of this as a crude form of probabilistic calibration; more could be done here.

They do not do any checking of individual-level predictions, though I don’t see why not. The model gives the probability for *each* individual who is unemployed to find work each week, so it would be straightforward to say something about the accuracy of those predictions. One could check calibration graphically, by Brier score, by relative entropy, etc.

Critiques Some have already been mentioned. Two more seem worth pointing out.

1. Why logistic GLM rather than an additive model with a logistic link function? Or a prediction tree? Or... well, you get the picture. Econometricians are addicted to linear models, and have pushed them on other social scientists, but it is really no harder to type `gam` than `glm`, and there is no good reason any more not to do so. Using more flexible statistical models *might* not alter the substantive conclusions, but why not check?
2. It’d be good to have of geographic spread of individuals’ social networks and how closely their network-neighborhood unemployment tracks their geographic neighborhood’s; this may in fact depend on covariates. (E.g., it would not be surprising if low-education immigrants had more geographically-localized networks than high-education natives.)

Accomplishment Despite my kvetching, they’ve done something good here. This is an “empirically-calibrated agent-based model” of an important real-world phenomenon, with sane models and results.

References

- Elster, Jon (1982). “Marxism, Functionalism and Game Theory: The Case for Methodological Individualism.” *Theory and Society*, **11**: 453–482. URL <http://www.jstor.org/pss/657101>.
- (1985). *Making Sense of Marx*. Cambridge, England: Cambridge University Press.
- Hedström, Peter (2005). *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge, England: Cambridge University Press.
- Kautsky, Karl (1988). *The Materialist Conception of History*. New Haven, Connecticut: Yale University Press. Abridged, annotated, and introduced by John H. Kautsky, translated by Raymond Meyer with John H. Kautsky from *Materialistische Geschichtsauffassung*.
- Oomes, Nienke A. (2003). “Emerging Markets and Persistent Inequality in a Nonlinear Voter Model.” In *New Constructions in Cellular Automata* (David Griffeath and Christopher Moore, eds.), pp. 261–284. Oxford: Oxford University Press.
- Shalizi, Cosma Rohilla and Christopher Moore (2003). “What Is a Macrostate? From Subjective Measurements to Objective Dynamics.” *Studies in History and Philosophy of Modern Physics*, **submitted**. URL <http://arxiv.org/abs/cond-mat/0303625>.