Unpacking Social Interactions
Ethan Cohen-Cole\textsuperscript{1}
and
Giulio Zanella\textsuperscript{2}

JEL Classification Codes: Z13
Keywords: Social Interactions, Conflation Problem

Abstract
As empirical work in identifying social effects becomes more prevalent, researchers are beginning to struggle with identifying the composition of social interactions within any given reference group. In this paper, we present a simple econometric methodology for the separate identification of multiple social interactions. The setting under which we achieve separation is special, but is likely to be appropriate in many applications.

Ethan Cohen-Cole
Federal Reserve Bank of Boston
600 Atlantic Avenue
Boston, MA 02210
(617) 973-3294
Email: ethan.cohen-cole@bos.frb.org

Giulio Zanella
University of Siena
Department of Economics
P.za S. Francesco 7
53100 Siena Italy
Email: zanella@unisi.it
I. INTRODUCTION

Following Manski’s (1993) seminal contribution to the econometric literature on social interactions, as well as Brock and Durlauf’s (2001) comprehensive treatment, researchers have taken seriously the identification problems that surround the empirical detection of social influences in individual behavior. To date, virtually any empirical work in this field addresses the so called “reflection problem”, using either econometric or experimental strategies.

A possibly trickier problem arises from the consideration that identification of social influences says nothing about their source. Suppose that one identifies correctly the causal effect of the prevalence of a certain behavior – such as crime, work effort, labor supply, human capital investment, welfare use, risky sexual behavior, unhealthy habits, etc. – in a reference group on individual behavior within that group. One cannot say whether the social effect is due to information sharing (people communicate and pass along information that increase the likelihood of behaving in a certain way), social learning (people observe others’ behavior and infer the distribution of outcomes from taking certain actions), stigmatization (the prevalence of a certain behavior makes its adoption less embarrassing), identity (a certain behavior is the hallmark of membership and loyalty to a group characterized by a precise identity), pure desire for imitation (keeping up with the Joneses), and so on. More likely, the estimated social effect will be the compound effect of all or some of these. Separate identification of different social effects, i.e. of the channel through which social interactions operate, is only apparently of second order importance compared to identification toutcourt. For instance, the aforementioned effects do not work necessarily in the same direction, and may offset each other. Therefore, in principle it is possible that estimated social effects are negligible while underlying social interactions do affect behavior.
Albeit extreme, this possibility suggests the issue is a relevant one. In particular, the separate identification of different social effects is important for policy evaluation because different channels of influence generally require different policies. Consider for example the literature on why "welfare cultures" emerge, i.e. which social factors affect welfare participation. Information sharing and stigma are natural candidates—and in fact have received considerable attention in both research and policy circles. Suppose a policy-maker wants welfare programs to reach all individuals they are designed for, and wants to take actions in order to increase take-up rates (which are notoriously low in the US). The appropriate measures depend on the relative magnitude of stigma and information effects. If stigma is the main cause of nonparticipation, then one could think of ways to “hide” recipiency—such as the replacement of food stamps with plastic cards. If instead information sharing is the main cause then one should think of ways to inform eligible nonparticipants. The point is that policies that work if one kind of social effect is predominant may be ineffective if instead another kind is predominant. However, even the most recent empirical works that identify social interactions in welfare participation such as Bertrand et al. (2000) do not attempt to identify them separately. An estimate of "aggregate" social interactions, despite important, may be of limited use because without knowledge of the composition and/or source of the social effect, one is hard pressed to make policy recommendations. The issue, which we label the conflation problem\(^1\), was brought to light by Manski (2000), and is a pressing one for empirical study of social interactions.

In this paper we suggest a simple and easily usable procedure to separately identify different kind of social interactions.\(^2\) We manage separation of a compound social interactions effect by exploiting the possibility that different reference groups are associated with different social effects. The procedure allows for overlapping reference groups, as well as non-exclusive
social effects, in the sense that two or more groups can be the source of the same effect. Effectively, we introduce an additional piece of sociological information with respect to existing work—that certain types of social interactions can be related with particular associations that a person might have. For instance, one might imagine that an individual shares information with his/her colleagues, but forms expectations about the outcome of certain actions (social learning) only considering the outcomes of individuals similar to himself/herself, and suffers social stigma from colleagues as well as neighbors and family. However, one does not need to interpret "distinct groups" in a special sense. The procedure we suggest works if groups are spatially or temporally distinct, or both.

Our result is based on the fact that external information on the source of interactions can provide a basis for disentangling a compound effect. As such it is based on prior information. Sometimes the researcher has particular prior information that allows the discrimination among different social effects, or to rule out a particular channel, in a simpler way than we suggest. For instance, Mas and Moretti (2006) identify peer influences on individual productivity among cashiers at a large supermarket chain. Then they ingeniously use information on spatial orientation and location of registers to conclude that the predominant social effect is social pressure—that is, workers monitoring each other. In another context, Aizer and Currie (2004) identify network effects in the use of publicly funded maternity services, based on panel data on first and second pregnancies. Then they devise a simple test arguing that if information is to be the predominant social effect then the magnitude of total social effects must be smaller for second deliveries. In the absence of particular information of this kind, one has to justify patterns of social influence on the basis of ethnographic (see Iyer and Weeks 2004) or sociological (see Cohen-Cole and Zanella 2006) evidence.
In the next section we outline our methodology in the context of a classic linear-in-means model of social interactions, and Section 3 concludes.

II. ECONOMETRIC MODEL

In this model, we allow individual to have multiple reference groups, a tool by which we will be able to distinguish different types of social effects. We consider the simple case of two social effects, labeled A and B, and two reference groups, labeled g and h. This case is not only analytically convenient, but is also frequently encountered in applications. For instance, preferences may be influenced by one group via social norms, but expectations by another group via social learning, as in Iyer and Weeks’ (2004) study of multi-level interactions in fertility. In this case our separation procedure works under two conditions: (i) the researcher can identify as many distinct reference groups influencing individual behavior as the number of social effects to be identified, in this case two, and (ii) each group is the source of at least one distinct effect, but not all effects. For reasons to be clarified below, we will also require that when the same social effect comes from different groups, the researcher has prior information on the importance of each such group as a possible source of social interactions of that given type.

Individuals choose an action $\omega_i$. We assume that individuals whose reference groups are $g$ and $h$ respond to the actions of other individuals in both groups, but asymmetrically. Suppose that social interactions of type B occur within both groups, but that interactions of type A occur within group g only. Under these assumptions, the canonical linear-in-means model posits:

\[
(1) \quad \omega_i = aX_i + bY_g + dY_h + J_g^{AB} m_g + J_h^{B} m_h + \epsilon_{igh}
\]

where $\omega_i$ is the individual choice under study, $X_i$ is a vector of individual controls (which include a constant), $Y_g$ is a vector of contextual controls common to all members of group $g$, $Y_h$ is a vector of contextual controls common to all members of group $h$, $m_g$ is mean choice in group $g$,
and $m_h$ is mean choice in group $h$. Using terminology from Manski (1993), $a$ expresses individual effects, $b$ and $d$ contextual effects, and $J_g^{AB}$ and $J_h^B$ endogenous social effects. Assume that the model is identified\textsuperscript{vi}, i.e. we have estimated $\left(a, b, d, J_g^{AB}, J_h^B\right)$. Clearly we have not identified, say, the magnitude of effects $B$ from group $g$, since there are infinite combinations $\left(J_g^A, J_g^B\right)$ such that $J_g^A + J_g^B = J_g^{AB}$. This is the essence of the conflation problem: the estimated values of $J_g^A$ and $J_h^B$, i.e. the effects of interactions of type A and B within group $g$, conflate into the compound estimate $\bar{J}_g^{AB}$. This is a problem because the researcher might be interested in estimating the "detailed" vector $\left(a, b, d, J_g^A, J_g^B, J_h^B\right)$.

The problem can be solved as follows. Define a "total effect $B$ function", which captures the effect of social interactions of type $B$ only in any group, $\phi \left(m_g, m_h\right)$. Therefore, since this function controls for effect $B$, we can write another model in which, by construction, mean behavior in group $g$ captures the effect of interactions of type $A$ only:

\begin{equation}
\omega_i = \alpha X_i + \beta Y_g + \delta Y_h + J_g^A m_g + J_h^B m_h + \epsilon_{gh}
\end{equation}

Estimating models (1) and (2), which we label respectively primary and auxiliary, it is possible to recover the detailed set of parameters as well as standard errors. First, let’s specify the $\phi(\cdot)$ function as linear:

\begin{equation}
\phi \left(m_g, m_h\right) = J_{gh}^B \left(\theta_g m_g + \theta_h m_h\right)
\end{equation}

where $J_{gh}^B$ has the structural interpretation of total effect $B$ from both groups $g$ and $h$. The weights $\theta_g$ and $\theta_h$ are assumed known, and represent the importance of groups $g$ and $h$ for interactions of type $B$. Such weights could be, for instance, demographic weights\textsuperscript{vii}. The
linearity assumption may appear arbitrary, and in fact, there is nothing that mandates that the "total effect $B$ function" be a linear relationship. The solution, however, is more tractable in the linear form. To summarize, by construction the primary model includes $g$-group effect for social reasons $A$, $B$ and the $h$-group effect for social reason $B$. The auxiliary model includes the $g$-group effect for social reasons $A$ alone and the joint $g$- and $h$-group effect for social reason $B$ alone. At this point, it may be useful to discuss an example underlying the social "reasons" above. As mentioned, the idea here is to use information on agent’s response to distinct groups’ average actions to unpack the composition of total social effects. Consider the case of student achievement, where $\omega$ measures test scores. Allow $g$ to refer to boys and $h$ girls. Now we have a common linear-in-means model with the exception that we have two endogenous coefficients to estimate.viii In many studies, readers are asked to assume that students respond to the efforts of other students in their decision to work hard themselves. Imagine, in this case that students respond both to the efforts of other students ($A$), but also to the social stigma associated with being studious ($B$). ix To estimate this effect, imagine that boys (girls) have knowledge of the amount of work done by other boys (girls) in their classroom, but don’t know about the study habits of the other gender. However, the social penalty is imposed by all students in the class. Thus, from the perspective of a boy, we can interpret the primary equation (1) as estimating the joint social effect for the boy of both study time and social penalty ($J_{g}^{B}$) and the social penalty from the girls ($J_{g}^{B}$). The auxiliary equation (2), coupled with the linearity assumption (3), on the other hand estimates the effect of study time alone from boys in the class ($J_{g}^{A}$) and the joint social penalty
from both genders \( J^B_{gh} \). In this example, one could potentially use the proportion of students of each gender to weight the effects \( \theta_g \) and \( \theta_h \) in the auxiliary regression. In this fashion, will be able to evaluate the degree to which students consider each of the two effects \( A, B \).

Since \( \theta_g \) and \( \theta_h \) are constant, no new information information is used in the auxiliary model with respect to the primary model. As a consequence we can state the following:

**Proposition 1:** Under the functional form (3), models (1) and (2) are informationally equivalent, in the sense that the OLS residuals are the same, \( \hat{\eta}_{igh} = \hat{\varepsilon}_{igh} \), and so are the estimated individual and contextual effects, \( (a, b, d) = (\alpha, \beta, \delta) \).

**Proof of Proposition 1:** Define two vectors of regressors:

\[
\begin{align*}
z'_{1i} &= (X_i, Y_g, Y_h, m_g, m_h), \\
z'_{2i} &= (X_i, Y_g, Y_h, \theta_g m_g + \theta_h m_h)
\end{align*}
\]

so that

\[
(4) \quad z'_{2i} = (I - A) z'_{1i}
\]

where

\[
A = \begin{bmatrix}
0 & \cdots & 0 & 0 \\
\vdots & & \vdots & \\
0 & \cdots & -\theta_g & 1-\theta_h
\end{bmatrix}.
\]

Rewrite primary and auxiliary models in matrix regression form using (4):

\[
(5) \quad \Omega = Z_1 \gamma + \varepsilon
\]

\[
(6) \quad \Omega = Z_2 \psi + \eta = Z_1 (I - A)' \psi + \eta
\]

where \( \gamma \equiv (a, b, d, J^A g, J^B h)' \) and \( \psi \equiv (a, b, d, J^A g, J^B h)' \). The OLS estimators are:

\[
\hat{\gamma} = (Z'_1 Z_1)^{-1} Z'_1 \Omega
\]
\[ \tilde{\psi} = (Z'_2 Z'_2)^{-1} Z'_2 \Pi = \left( (I - A) Z'_1 (I - A)' \right)^{-1} Z'_1 (I - A)' \Pi = \left[ (I - A)' \right]^{-1} (Z'_1 Z_1)^{-1} Z_1 \Pi. \]

Hence,

\[ (7) \tilde{\psi} = \left[ (I - A)' \right]^{-1} \hat{\gamma}. \]

Therefore, using (5), (6), and (7):

\[ \hat{\eta} \equiv \Omega - Z_1 (I - A)' \tilde{\psi} = \Omega - Z_1 (I - A)' \left[ (I - A)' \right]^{-1} \hat{\gamma} \equiv \hat{\epsilon}. \]

Furthermore, (7) reveals that \( (\bar{a}, b, d) = (\bar{a}, \beta, \delta) \). QED.

A corollary of Proposition 1 is that when one of the two models is assumed to be the "true model," the other one can be defined as true as well. This in turn implies that coefficients are comparable across the primary and auxiliary model. Furthermore, we can rewrite the latter as follows:

\[ \omega = aX_i + bY_g + dY_h + J^A m_g + J^B \left( \theta_g m_g + \theta_h m_h \right) + \varepsilon_{gh}. \]

The comparability of coefficients across models (1) and (8) suggests a simple way of estimating the marginal effect of interest that, after identifying primary and auxiliary models, is still conflated with the others, namely \( J^B_g \). We suggest the following estimator:

\[ (9) J^B_g \equiv J_{gh}^B - J_h^B. \]

This makes intuitive sense: since under our assumptions we can compare coefficients across models, to obtain the effect of social interactions of type B from group g only, we subtract from the total effect the portion that does not come from group g. By the same token, one could also write

\[ (10) \bar{J}^B_g \equiv \bar{J}^B_g - \bar{J}^A_g. \]
where one simply subtracts the social effect $A$ from the compound effect within group $g$. Of course these two estimators are equivalent under our assumptions. To see this, compute marginal effects from the primary and auxiliary regression equations:

$\frac{\partial \omega_i}{\partial m_h} = J^B_h = \theta_h J^B_{gh}$,  $\frac{\partial \omega_i}{\partial m_g} = J^{AB}_g = J^A_g + \theta_g J^B_{gh}$

Replace the first of these equations into (9):

(11) $\hat{J}^B_g = \hat{J}^B_{gh} - \hat{\theta}_g \hat{J}^B_{gh} = \theta_g \hat{J}^B_{gh}$,

and the second into (10):

(12) $\hat{J}^A_g = \hat{J}^A_g + \hat{\theta}_g \hat{J}^B_{gh} - \hat{J}^A_g = \theta_g \hat{J}^B_{gh}$

whence

(13) $\hat{\sigma}_{J^B_g} = \theta_g \hat{\sigma}_{J^B_{gh}}$

This provides a way to estimate, $J^B_g$, and its standard error unequivocally, along with a useful specification test: if the model is correctly specified, then estimators (9) and (10) should provide the same estimate. Also, $\hat{J}^B_g$ inherits all the properties of $\hat{J}^B_{gh}$, notably consistency and asymptotic normality. However, while our estimators are statistically well-behaved, their structural interpretation should be regarded as resting on an approximation: in estimating $J^B_{gh}$ through model (8) it is assumed, by linearity of $\phi(m_g, m_h)$, that social interactions of type $B$ in groups $g$ and $h$ are, so to speak, perfect substitutes, a restriction that is not imposed when estimating $J^B_h$ through model (1). Therefore, (9) compares two estimators that incorporate two possibly conflicting views on the sensitivity of individuals to social effects from different groups.

This possible logical inconsistency requires regarding $\hat{J}^B_g$ as approximating the effect of $B$ from
group $g$, without corresponding measures of precision; i.e. we have little understanding of the quality of this particular approximation. Of course this holds for the entire model. Indeed equation (7) also establishes the following relations between social interactions parameters:

\[(14) J^B_h = \theta_g J^B_g,\]

\[(15) J^A_g = J^A_g - \left(\frac{\theta_g}{\theta_h}\right) J^B_h.\]

which are subject to the same caveat. Also notice that equation (14) mimics (11) and (12), which suggests our estimator is consistent with the structure of the model.

III. CONCLUSION

In this paper we have argued that when trying to estimate the effect of social interactions on economic behavior, one needs to address explicitly the fact that different social effects are possibly at work. As such, we presented a simple econometric method for the isolation of distinct social interactions in the linear-in-means model. This method is appropriate when the researcher has prior knowledge that a given group is a source of two social effects, while a different--possibly partially overlapping--group is a source of only one of the two. Our opinion is that this method is well suited to contexts in which there exists prior qualitative research that indicates distinction in social effects at different layers of society--though our guess is that there are many such occasions.

References


**Notes**

---

i Federal Reserve Bank of Boston 600 Atlantic Avenue Boston, MA 02210. (617) 973-3294 Email: ethan.cohen-cole@bos.frb.org. We are grateful to Hisa Tanaka for very useful comments, as well as to Nick Kraninger for research assistance. The views expressed in this paper are solely those of the authors and do not reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System.

ii University of Siena, Department of Economics P.za S. Francesco 7, 53100 Siena Italy. Email: zanella@unisi.it.

1 We wish to thank Giacomo Rondina for suggesting this term to us.

2 In a companion paper (Cohen-Cole and Zanella, 2006) we have applied the procedure to estimate the magnitude of information and stigma effects in welfare participation, using US data.

iii This could be generalized to more than two social effects, which is something we have not yet attempted.
As is well known, one’s measure of social interactions is only valid conditional on a correct definition of social reference groups. If agents do not respond to the mean behavior of agents in $g$ and $h$, then the estimates of social interactions may be severely biased.

In a companion paper to this one (Cohen-Cole and Zanella, 2006), we hypothesize that the choice to participate in welfare is mediated by information and stigma effects from one’s own race-ethnic group at the local level but stigma effects alone from the remainder of the community.

Identification is a concern because mean behaviors are endogenous, i.e. $m_g = E(\alpha_t \mid Y_g)$ and $m_h = E(\alpha_t \mid Y_h)$, which is the source of the reflection problem. Effectively, in order to focus on separate identification, we assume this and other inferential problems (see Brock and Durlauf, 2001) have been solved.

This expedient was used by Bertrand et al. (2000) to achieve identification, and is sensible if weights are chosen in a meaningful way.

Others have estimated more than one coefficient in a social interactions context, though for varying purposes. See for example Conley and Topa (2002).

It is well known that in many American schools, there is a social penalty of sorts for students to perform too well academically. See Akerlof and Kranton (2002) for a theoretical model.