Social Insurance Networks

By

Markusen, Simen og Knut Røed

This is a post-peer-review, pre-copyedit version of an article published in:

Journal of Human Resources

The definitive publisher-authenticated and formatted version:

Markusen, Simen og Knut Røed, 2015, Social Insurance Networks, Journal of Human Resources, Vol 50(4), 1081-1113, DOI: 10.3368/jhr.50.4.1081.

is available at:

https://doi.org/10.3368/jhr.50.4.1081

August, 2014

Social Insurance Networks

Simen Markussen and Knut Røed*

The Ragnar Frisch Centre for Economic Research

ABSTRACT

Based on administrative panel data from Norway, we examine how social insurance claims spread among neighbors and former schoolmates. We use a fixed effects methodology that accounts for endogenous group formation, contextual interactions generated by predetermined social factors, and time-constant as well as time-varying confounders. We report evidence that social insurance claims are contagious. There are significant local peer effects both in the overall use of social insurance and in the propensity to use one particular social insurance program rather than another. The magnitudes of the estimated peer effects rise consistently with measures of geographical and relational closeness.

Keywords: social interaction, social multiplier, work norms, peer effects.

JEL Classification: C31, H55, I38,

* This paper is part of the project "Social Insurance and Labor Market Inclusion in Norway", funded by the Norwegian Research Council (grant #202513). Data made available by Statistics Norway have been essential for the research project. Thanks to Bernt Bratsberg, Oddbjørn Raaum, three anonymous referees, and the Editor for valuable comments. Correspondence to: Knut Røed, the Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, 0349 Oslo, Norway. E-mail: knut.roed@frisch.uio.no.

I. Introduction

The purpose of this paper is to examine peer effects in social insurance (SI) claims. The paper is motivated by two observations. First, there has been a conspicuous – yet basically unexplained – rise in social security dependency in many countries, particularly related to health problems; see, e.g., Duggan and Imberman (2006), Bratsberg *et al.* (2013), and Burkhauser and Daly (2011). And second, there tend to be correspondingly large and unexplained geographical disparities in dependency rates as well as in attitudes toward social insurance both within and across countries; see McCoy *et al.* (1994), OECD (2010), and Eugster *et al.* (2011). Although far from offering a complete explanation, these empirical patterns may be easier to understand if SI claim propensities exhibit path-dependency due to peer effects; see, e.g., Bertrand *et al.* (2000) and Durlauf (2004). Such peer effects could result from transmission of work norms or changes in the stigma attached social insurance claims (Moffitt, 1983; Lindbeck, 1995; Lindbeck *et al.*, 1999; 2003), or they could arise from the transfer of information about eligibility rules, application procedures, and acceptance probabilities (Aizer and Currie, 2004), or about job opportunities (Ioannides and Loury, 2004).

While social interaction effects have been extensively analyzed from a theoretical perspective, empirical analysis has been held back by methodological difficulties and lack of appropriate data. The fundamental empirical challenge is to disentangle endogenous interaction from other sources of correlation between individual and group behavior, such as endogenous group formation and unobserved confounders; see Manski (1993). As shown by our brief literature review in the next section, the existing empirical evidence on SI contagion is scant and, with a few important exceptions, limited to ethnic minorities. Existing evidence is also confined to very specific SI programs, making it difficult to assess whether it has captured peer effects in *overall* SI dependency or in the tendency to use one specific SI program rather than another. The policy implications following from these two competing interpretations are clearly different.

In the present paper, we examine social interaction effects within different kinds of networks – or peer groups – i.e., neighbors, schoolmates, and ethnic minorities. The key research question we ask is whether – and to what extent – an agent's likelihood of claiming *any form of tax-financed income support* is causally affected by the level of claims recorded within the various types of networks the agent relates to, conditional on the claim patterns prevailing elsewhere in the economy. In addition, we examine peer effects in the propensity to use one type of SI program rather than another. The question of interest here is whether – and to what extent – the distribution of SI claims between "disability-related" and "unemployment-related" programs within a network affect the group members' propensities to claim benefits from these program types.

We use an extraordinarily rich and detailed panel data set from Norway, covering the whole working-age population over age 17. We exploit the richness of the data to set up empirical models in which we control for the various confounding and sorting problems that often undermine the credibility of reported social interaction effects. In contrast to much of the existing literature, we do not rely on either instrumental variables or movements between networks, but instead use individual fixed effects to remove the influence of time-constant confounders and contextual interactions generated by predetermined social factors, and flexible time functions to control for network-specific shocks and sorting problems that are not eliminated by the individual fixed effects. A novel feature of our empirical approach is that we examine how SI interaction effects vary with *geographical as well as relational distance,* i.e., we are not only interested in effects of peer-group behavior per se, but also in the way the interaction effects vary as we move from "close" to more "distant" network members.

Our findings confirm the empirical relevance of social interaction. We present several empirical results indicating that individuals' own SI claim propensities are significantly affected by claim patterns among peers, and that the effects grow with relational closeness. For example, we find that an exogenous change in average SI claims within a group of adults who at some time went to the same junior high school together generates cumulative knock-on effects amounting to around 25% of the initial change. But, adjusted for group size, the peer effect is *much* larger among same-level-same-sex schoolmates than it is among more "distant" schoolmates. Within small neighborhoods, we find that an exogenous change in SI claims entails additional knock-on effects amounting to 17% of the initial change. Again, the effect is much larger among similar than among dissimilar neighbors, and also larger among geographically close than among geographically more distant neighbors. We find particularly strong interaction effects within ethnic networks, defined as immigrants from a common low-income source country who reside in the same local area. The cumulative peer effect for these groups is estimated to around 38%. The peer effects do not cross ethnic boundaries, however; a rise in SI dependency among immigrants from *other* low-income countries in the same local area has no effect at all.

Our results also indicate considerable scope for substitution between different SI programs. When we distinguish between "disability-related" and "unemployment-related" SI claims, we find that an exogenous rise in a peer group's use of one of these program types has a much larger effect on group member's same-type-claims than it has on their overall use of SI. A significant part of the additional SI claims is thus offset by a *reduction* in claims of the other type. An important implication of this finding is that empirical approaches focusing on a single program only will tend to exaggerate the peer effects in overall SI dependency. Peer effects are important both for the overall level of SI claims and for the allocation of claims across SI programs.

Finally, we show that peers affect SI claim propensities among previous non-claimants (the entry decisions) as well as among more experienced claimants (the continuation decisions). This suggests that the peer-effects not only mirror a process of information-sharing, but also impacts on the utility associated with being in a state of benefit recipiency. And while peers' use of *competing* (other-type) SI programs has a negative (substitution) effect on *entry* into each program type, it has a positive impact on *continuation*. We conclude from this that the peer effects identified in this paper to a large extent reflect the propagation of work-norms or the stigma associated with being an SI claimant.

II. Related literature

There is by now a large and rapidly expanding empirical literature on social interactions within economics, covering a wide range of topics; see, e.g. Durlauf (2004) or Ioannides and Loury (2004) for recent reviews and Blume *et al.* (2010) for a comprehensive overview of the various identification strategies that have been applied in the literature. The latter paper concludes that the current research frontier still involves efforts to achieve identification in the presence of the three challenges originally highlighted by Manski (1993): i) to differentiate between social interactions that derive from direct interdependencies between choices (endogenous interactions) and social interactions that derive from predetermined social factors (contextual interactions), ii) to deal with the presence of group-level unobserved heterogeneity (confounding factors), and iii) to deal with the presence of endogenous formation of the groups that act as carriers of social interactions.

There is also a growing empirical literature on peer-effects in the utilization of public transfers. Bertrand *et al.* (2000) examine the role of welfare participation within local networks in the U.S., defined by language spoken. Their empirical strategy is to investigate whether belonging to a language group with high welfare use has larger effects on own welfare use the more a person is surrounded by people speaking one's own language. They find that this is indeed the case, and conclude that networks are important for welfare participation. Aizer and Currie (2004) use a similar approach to study network effects in the utilization of publicly funded prenatal care in California, with groups defined by race/ethnicity and neighborhoods. They conclude that group behavior does affect individual behavior. Furthermore, they show that the identified network effects cannot be explained by information-sharing, since the effects persist even for women who had used the program before. Conley and Topa (2002) examine the spatial patterns of unemployment in Chicago, and find that local variations are consistent with network effects operating along the dimensions of race and geographical and occupational proximity.

The recent literature also includes a number of studies outside the U.S. Stutzer and Lalive (2004) examine the pattern of unemployment duration in Switzerland, and find that strong local work norms – as measured by voting behavior in a referendum on the level of unemployment insurance – tend to coincide with short unemployment durations. Hesselius *et al.* (2009) use experimental data from Sweden to examine the extent to which co-workers affect each other's use of sick-pay. The experiment they use implied that a randomly selected group of workers were subject to more liberal rules regarding the need for obtaining a physician's certificate to prove that their absence from work was really caused by sickness. Hesselius *et al.* (2009) show that the reform caused absenteeism to rise both among the treated and the non-treated workers, and that the latter effect was larger the larger was the fraction of treated workers at the workplace. Peer effects in absenteeism are also examined by Ichino and Maggi (2000). Their empirical strategy is to study how workers who move between branches in a large Italian bank adapt to the prevailing absence cultures in the destination branches. The key finding is that workers adjust own absence behavior in response to the absence level among their new colleagues. A similar approach has been used by Bradley *et al.* (2007) to study absenteeism among school teachers in Queensland, Australia. And again, the finding is that the absenteeism of movers to some extent adapts to the prevailing absence culture at their new school. Åslund and Fredriksson (2009) examine peer effects in welfare use among refugees in Sweden, exploiting a refugee placement policy which generates the rarity of exogenous variation in peer group composition. A key finding of the paper is that long-term welfare dependency among refugees is indeed higher the more welfare-dependent the community is in the first place.

Empirical evidence on peer effects in the utilization of social insurance in Norway is provided in three recent papers by Rege *et al.* (2012), Bratberg *et al.* (2012), and Dahl *et al.* (2013), respectively. Rege *et al.* (2012) investigate neighborhood peer effects in disability insurance program participation among older workers by means of an instrumental variables strategy. Their key idea is that since the probability of disability program entry in Norway has been shown to be strongly affected by job loss (Rege *et al.*. 2009; Bratsberg *et al.*, 2013), exogenous layoffs in a person's neighborhood, e.g., caused by firm closure, can be used to instrument the neighbors' disability program participation (with proper controls for local variations in labor demand). Based on this strategy, Rege *et al.* (2012) estimate a sizable network effect implying that a 1 percentage point exogenous increase in similarly aged neighbors' disability program participation rate generates an additional increase of 0.3-0.4 percentage points as a result of network effects. Bratberg *et al.* (2012) and Dahl *et al.* (2013) both assess the transmission of disability pension recipiency within families, but with completely different empirical strategies. Bratberg *et al.* (2012) take the view that the intergenerational transmission of, say, work norms, operates through "exposure", and identify the social interaction effect by comparing siblings who, due to differences in age, to varying extent shared a household with their parent after the disability pension was granted. Their finding confirms that longer exposure to a parent claiming disability insurance indeed raises the probability that the offspring also claims such benefits later on. Dahl *et al.* (2013), on the other hand, mainly focus on children who were adults at the time of a parent's potential entry to the disability insurance program, and use a random assignment component in the decision process – the assignment of judges to applicants whose cases were initially denied – as the source of identification of the intergenerational transmission mechanism. Again, the key finding is that a parent's entry to the disability insurance program significantly raises the probability that their offspring also enter the program.

What all the pieces of Norwegian evidence have in common is that they mostly focus on the transmission of permanent disability insurance claims. We will argue that this is an unfortunate limitation, since permanent disability insurance typically either substitutes for or is preceded by other social insurance programs, such as sick pay, temporary disability insurance (medical or vocational rehabilitation benefits), unemployment benefits, or social assistance (welfare). A considerable degree of substitutability between unemployment and disability insurances has been established in several empirical papers; see, e.g., Black *et al.* (2009), Autor and Duggan (2003), Rege *et al.* (2009) and Bratsberg *et al.* (2013). Peer effects may thus be relevant both for the overall SI claim propensity and for the distribution of claims across programs. By focusing on a single program only, it is impossible to distinguish peer effects on the overall use of social insurance from peer effects on the propensity to use one particular program rather than another.

The present paper adds to the existing literature in at least two ways: The first is related to the substantive research questions: It is the first paper to address peer effects associated with all social insurance programs jointly. While our main interest lies in identifying the extent to which persons' propensities to claim social insurance benefits depend on the overall use of benefits among various peer groups, we also examine the impacts that peers' behavior have on the use of particular types of SI programs. Our second contribution relates to our empirical approach: In contrast to the existing literature, we build our identification strategy on the use of individual fixed effects; hence focusing on the observed *timing* of SI claims, rather than on their occurrence. This approach is backed up by the use of extraordinarily flexible control functions – with up to as much as 623,000 time-varying dummy variables – arguably eliminating the influence of conceivable time varying confounding factors.

III. Theoretical Considerations

Social interaction models start from the idea that the preferences of individuals over alternative courses of action depend directly on the actions taken by other individuals to whom the individuals relate; see, e.g., Brock and Durlauf (2000) and Cont and Löwe (2010) for overviews. The purpose of these models is typically to characterize or to provide an explanation for group behavior which emerges from interdependencies between individuals. To illustrate, let a_i indicate individual *i's* use of social insurance, and assume that the payoff function associated with this action can be decomposed into a sum of a private and a social component. Let a_i^0 denote the optimal choice in the absence of social interaction and let $j \in J$ be the set of agents that *i* relates to. With quadratic utility, we can write

$$U_i\left(a_i;\left\{a_j,\, j \neq i\right\}\right) = -\pi(a_i^0 - a_i)^2 - \sum_{j \neq i} \gamma_{ij}(a_i - a_j)^2, \tag{1}$$

with the optimal SI claim characterized by

$$a_i^* = \frac{1}{\pi + \sum_{j \neq i} \gamma_{ij}} \left(\pi a_i^0 + \sum_{j \neq i} \gamma_{ij} a_j \right). \tag{2}$$

In this specification, π reflects the marginal disutility of deviating from the private optimum and γ_{ij} measures the marginal gain in *i's* utility of conforming to the action of *j*. Note that it is the *actual behavior* of *j* that *i* conforms to, and not the norms/attitudes that motivate *j's* behavior; hence γ_{ij} represents what Manski (1993) refers to as *endogenous interaction*. While endogenous and contextual interactions both represent important social propagation mechanisms, it may be important from a policy perspective to discriminate between them, since only endogenous interactions are able to create spill-over or multiplier effects of policy interventions targeted at changing actual behavior. Formally, endogenous interactions imply that optimal choices are determined in a large simultaneous equations system, with as many equations as there are individuals.

We emphasize that SI contagion in this context does not necessarily reflect the prevalence of fraudulent claims. Both unemployment-related and disability-related SI programs involve substantial scope for subjective judgment with respect to whether or not a given health- or unemployment problem is sufficiently serious to justify benefits, and there is also considerable overlap between different programs. In addition, individuals can obviously exert more or less effort in order to prevent the need for SI to arise in the first place as well as to escape from it. Hence, even in cases with strict screening and monitoring (e.g., in the form of physician certification or job-search requirements), the utility (or disutility) associated with different types of benefit-recipiency matters for the realized level of claims.

Different classes of models are obtained from Equation (1) by parameterizing γ_{ij} in different ways. For example, the choice $\gamma_{ij} = \gamma / N$, where *N* is the size of the population (excluding *i*), leads to the global interaction model, where each agent's preferences are affected by the average action of all others, as in Lindbeck *et al.* (1999) and Glaeser *et al.* (2003). By contrast, local interaction models assume that social influences are mediated within confined groups, potentially differentiated by some notion of "distance" such that $\gamma_{ij} = \gamma(d_{ij})$, where d_{ij} is a measure of relational distance between *i* and *j*. Studies on the structure of social groups show that individuals tend to interact most with other individuals who are similar to themselves; see, e.g., Marsden (1982). In empirical applications, social interactions are thus typically assumed to take place within peer groups, defined in terms of, e.g., neighborhoods, workplaces, school-classes, families, or races, often in combination with demographic factors (gender, age) and measures of "social distance" (e.g., educational attainment or "class").

In the present paper, we focus on local interactions. Interaction effects are examined at group-levels, and group-averages are used as the central explanatory variables. This implies that the bivariate interaction effects – the direct influence of one person on another – are modeled as homogeneous within (narrowly defined) groups and inversely related to group size; i.e., $\gamma_{ij} = \gamma_g / N_g$, where *g* denotes the group in question and N_g is the number of group members apart from *i*. An important assumption embedded in this framework is that average distance increases with group size, ceteris paribus, such that the larger the number of peers in a particular group, the smaller is the influence exercised by each and one of them. Equation (2) can then be reformulated as

$$a_i^* = \frac{1}{\pi + \sum_g \gamma_g} \left(\pi a_i^0 + \sum_g \gamma_g \overline{a}_{g,-i} \right), , \qquad (3)$$

where γ_g is the utility of conforming to the average behavior in group *g* ($\overline{a}_{g,-i}$). This parameter clearly depends on the weight attributed by individual *i* to the behavior of group *g*, which is again a reflection of its relational closeness, and potentially also its size. We typically expect $\gamma_g \geq 0$, but $\gamma_g < 0$ can of course not be ruled out. Negative interaction effects may occur when agents derive utility from displaying novelty, as in fashion and fads, or from signaling a distance to groups one do not wish to be associated with.

IV. Institutional Setting and Data

The Norwegian public system of social insurance is comprehensive. In the present paper, we examine all the major social insurance programs relevant for the working age population in Norway; *i.e.*:

- Unemployment insurance
- Sick-pay (spells exceeding 16 days only)
- Temporary disability insurance (including medical and vocational rehabilitation)
- Permanent disability insurance
- Subsidized early retirement (starting at age 62)
- Social assistance (welfare)

Entitlement to unemployment insurance, sick-leave benefits and subsidized early retirement is obtained through regular employment, whereas rehabilitation benefits, disability pension, and social assistance in principle can be obtained without such experience. The replacement ratios for unemployment insurance, temporary and permanent disability, and subsidized early retirement all typically lie around 60-65% of previous earnings, but with minimum and maximum levels. For sick-leave, the replacement ratio is 100%, but these benefits can only be maintained for one year (persons who are still unable to work after one year of sickness can apply for temporary or permanent disability benefits). All disability-related benefits, including sick-pay, need to be certified by a physician. Yet, in practice it has turned out to be difficult for physicians to "overrule" their clients' own judgments; see Markussen *et al.* (2013). Social assistance constitutes the last layer of social insurance and is primarily targeted at individuals with no other income sources. In contrast to the other benefits, it is means tested against family income.[1]

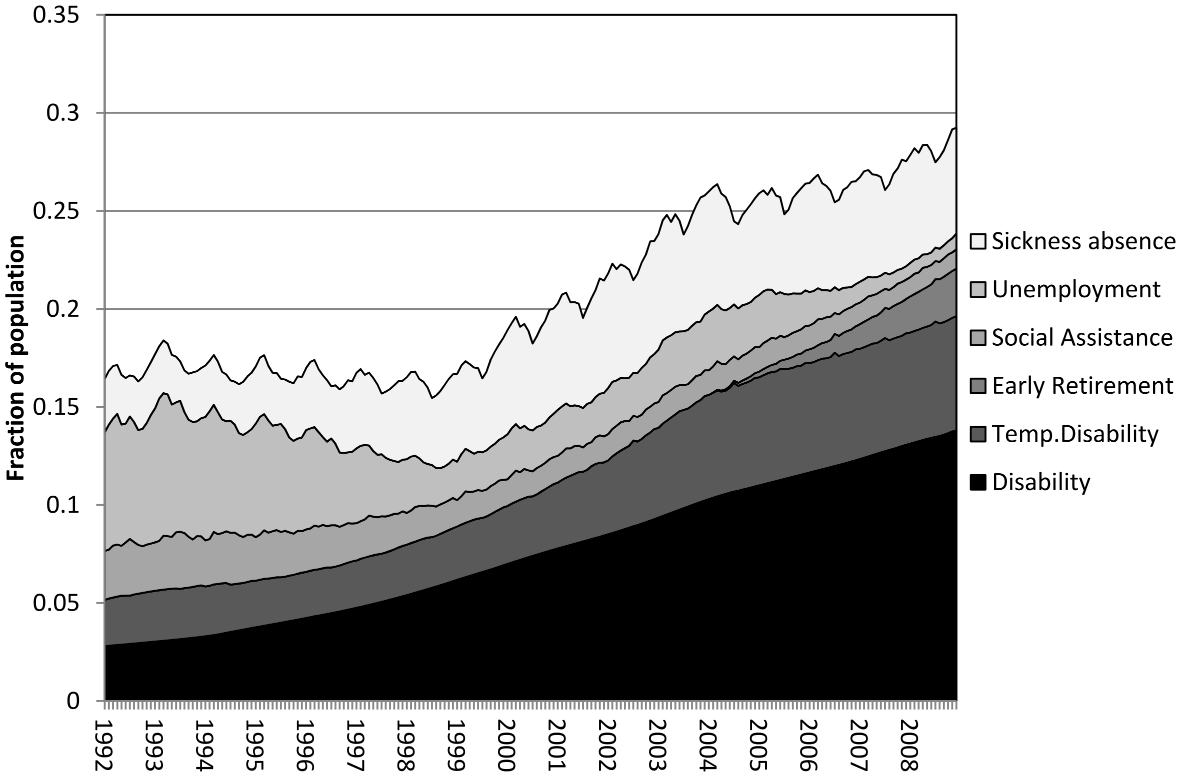

Figure 1. Social insurance claims for the 1942-1974 birth cohorts from 1992.1 to 2008.12
Note: Data include all persons who resided in Norway from 1992 to 2008 and who were born between 1942 and 1974 (1,867,662 individuals).

Our data cover social insurance claims for the whole Norwegian population from 1992 through 2008. Since we have chosen to use a balanced panel (see next section), we limit the analysis to individuals who were between 18 and 66 years throughout this period, implying that they were born between 1942 and 1974. This implies that our analysis comprises 33 complete birth cohorts, conditioned on being alive and residing in Norway in 1992-2008. Figure 1 gives an overview of these cohorts' social insurance claims – month by month – by SI program. Our primary interest does not lie in the use of each particular program, however, but

[1] Due to space considerations, we do not give a detailed description of Norwegian social insurance institutions here. More thorough descriptions (in English) are provided by Halvorsen and Stjernø (2008) and by the European Commission (2011).

rather in overall SI claims. This focus is partly motivated by the fact that the distinction between the different programs is blurred (Bratsberg *et al.*, 2013), with large flows between them (Fevang *et al.*, 2004), and partly by our ambition to identify patterns of interest beyond a narrow program-specific Norwegian setting. We are also not particularly interested in the high-frequency (month-to-month) fluctuations in SI use, which for some of the programs are dominated by seasonal factors. Hence, in the main part of our statistical analysis, we aggregate the observed social insurance outcomes into an annual dependent variable measuring the number of months with benefit claims from any of the social insurance programs in Norway.[2] However, to illuminate how peers potentially affect the selection of particular SI programs, we also set up models where we distinguish the presumed disability-related programs (sick-pay, temporary and permanent disability benefits, early retirement benefits) from the presumed unemployment-related programs (unemployment benefits, social assistance).

[2] For some of the programs, we are not able to identify accurately the dates, or the number of days, with benefit receipt; we only observe whether or not the benefit in question was received during each month.

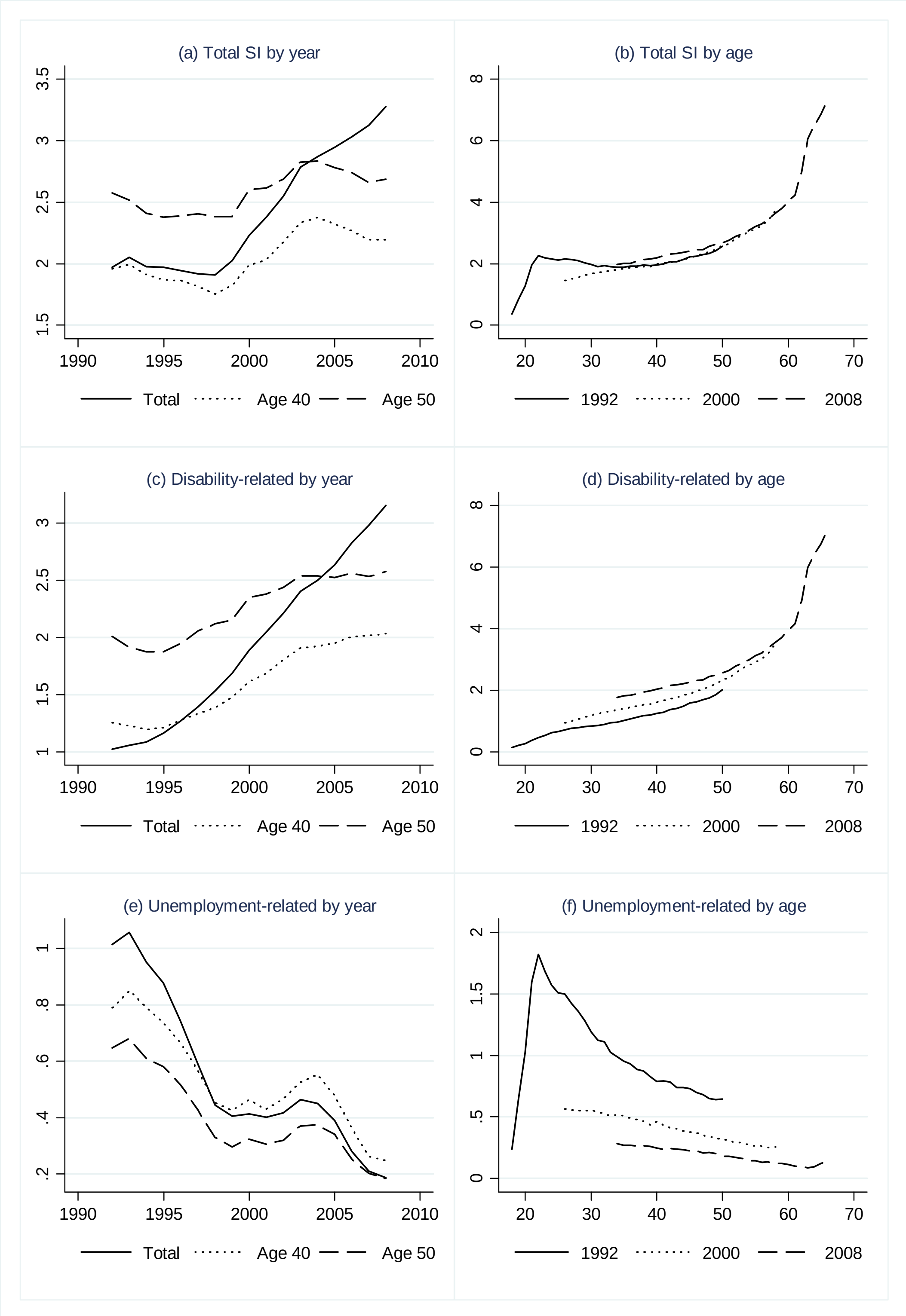

Figure 2. Number of months with SI claims in Norway, by year (1992-2008) and age (18-66).

Note: Data include all persons who resided in Norway from 1992 to 2008 and who were born between 1942 and 1974 (1,867,662 individuals).

Figure 2 illustrates some key descriptive features of the dependent variables that we are going to use in the empirical analysis. The two upper panels show how the overall use of SI developed within our analysis population from 1992 through 2008, by year and age, respectively. Since we follow the same group of people over time in this analysis, it is clear that the strong age gradient shown in panel (b) is an important factor behind the observed trend in social insurance claims shown in panel (a). This is also illustrated by the much weaker time-trend observed in panel (a) when the age-level is fixed (at age 40 and age 50, respectively). It still seems to be the case, though, that the overall SI caseloads rose significantly in the period from around 1993 to 2003, after which there was a small decline. The four lower panels illustrate the corresponding developments for disability-related and unemployment-related SI claims separately. They reveal a sharp increase in disability-related claims and a decline in unemployment-related claims.

The important role that age seems to play in the determination of individual SI claims suggests that the social interaction effects generated by a given average SI use among peers may depend on the age-composition of the peer group in question. For example, a high SI rate primarily caused by a large fraction of elderly individuals in the peer group may have a different impact on work morale than the same high rate caused by unusually high claimant rates among younger individuals. We will therefore use age-adjusted peer group averages in the statistical analyses; i.e., for each person-year in the peer group, we subtract the grand (national) age-specific mean for the year in question and then add the corresponding mean for 40-year-olds. As a result, we obtain age-adjusted observations normalized to a person aged 40.

During the period of declining unemployment and rising disability-related SI claims in the 1990's an interesting cross-sectional pattern emerged, whereby the local rises in disability-related claims tended to be larger the steeper were the declines in unemployment-related claims. This is illustrated in Figure 3, where we for 1,535 local areas in Norway (to be described in the next section) plot the changes in average age-adjusted disability claims from 1993 to 2003 against the corresponding changes in unemployment-related claims. The marked inverse relationship between these local trends raises the question of whether a causal relationship exists. One potential source of such a relationship could be program substitution generated by “cross-program” peer effects. For example, it is conceivable that in areas with a particularly sharp decline in unemployment due to a booming labor demand, it became less attractive to present a given “labor market problem” as being caused by unemployment and more attractive to present it as being caused by poor health.

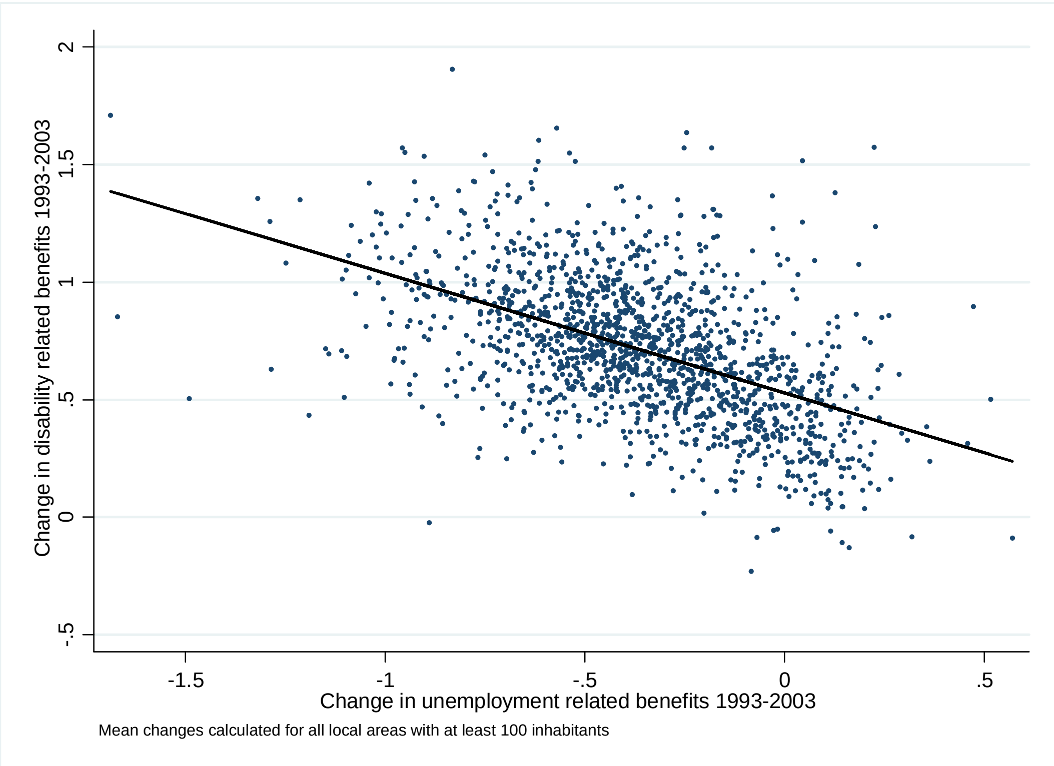

Figure 3. Changes in age-adjusted social insurance (SI) claims in 1,535 local areas in Norway.
Note: The regression line (OLS) has slope -0.50 (standard error 0.02)

V. Empirical Analysis

In this section, we set up linear regression models designed to find out whether – and to what extent – an individual’s use of social insurance benefits is causally affected by the (age-adjusted) use within networks/groups that the individual is closely – or more vaguely – attached to. Our primary dependent variable is going to be a person-year observation reporting the number of months with SI claims, either in total or for disability- and unemployment-related programs separately. The key explanatory variables are the following: i) a person-fixed effect, ii) the individual's own claim last year, iii) the average claims among peers last year, and iv) a vector of group-year fixed effects, where – essentially – the grouping does not coincide exactly with the peer groups. In some specifications, we also use observed time-varying variables defined at the peer-group level to account for possible confounders.

To circumvent the problem of *dynamic* endogenous group-formation, we focus throughout this paper on groups that – by definition – are stable; i.e., former schoolmates and persons that resided in the same geographical area at the start of our observation period. The price we pay for this is that our "networks" will serve as imperfect proxies for the various groups of people that agents actually interact with. Hence, compared to analyses based on positively identified and closely tied networks, we expect that interaction effects identified in our analysis will be significantly attenuated.

We first present the model for total SI claims. Let $y_{i,t}$ be the number of months that individual *i* claimed (any form of) SI benefit in year *t* and let $\overline{y}_{g,-i,t}$ be the corresponding age-adjusted SI propensity (see previous section) for persons belonging to a group *g* in year *t*, excluding individual *i*. We set up fixed effects models of the following form:

$$y_{i,t} = \alpha_i + \rho y_{i,t-1} + \lambda h_t\left(x_i\right) + \sum_{g \in G} \theta_g \overline{y}_{g,-i,t-1} + u_{it}, \tag{4}$$

where α_i is an individual fixed effect, $h_t\left(x_i\right)$ is a time function specified separately for different combinations of individual covariates x_i, and *G* is the set of groups/networks potentially influencing the behavior of *i*. The parameters of main interest are the θ_g's, which reflect the first-year peer effects. There will also be knock-on effects in subsequent years, as the first-year effect propagates both through the autoregressive process and through additional higher-order peer effects. The knock-on effects will decline over time provided that $\left(\theta_g + \rho\right) < 1$. Consider an exogenous and transitory shock in group *g's* SI dependency of size *z*. Next year, this shock implies a change in SI dependency equal to $\left(\theta_g + \rho\right) z$, and the year after that $\left(\theta_g + \rho\right)^2 z$ and so on. Hence, the cumulative knock-on effects arising from the shock amounts to $z\left[\left(\theta_g + \rho\right) + \left(\theta_g + \rho\right)^2 + \ldots\right]$, which converges toward $z(\theta_g + \rho)(1 - \theta_g - \rho)^{-1}$. Without peer effects, it would instead converge toward $z\rho(1-\rho)^{-1}$. Hence, as a measure of the total cumulative *peer effect*, we compute the statistic

$$\tau_g = \frac{\theta_g + \rho}{1 - \theta_g - \rho} - \frac{\rho}{1 - \rho}, \tag{5}$$

which represents the total number of extra SI months – over and above what arises from the autoregressive process alone – that accrues in response to a one-year transitory shock in the group average due to the peer effects.

The individual fixed effect (α_i) is included in Equation (4) to control for sorting on overall SI-propensity into networks, for time-constant confounders, and for predetermined contextual sources of interaction. It ensures that it is the *timing* – not the occurrence – of SI claims within networks that identifies the effects of interest At first sight, it may appear un-necessary to use individual fixed effects in this setting, since it is disturbing factors at the network level that we primarily worry about. However, without individual fixed effects, we could not have justified the critical assumption of exogenous peer behavior in year *t-1,* since it would have been affected by individual *i's* own (unaccounted for) SI propensity in earlier years.

The lagged dependent variable $(y_{i,t-1})$ is included to account for the strong autocorrelation present in SI claim patterns. If unaccounted for, this pattern will cause a simultaneity problem, since a person's past SI statuses (more than one year ago) will have had a causal impact on the peers' SI statuses last year and at the same time be correlated with the individuals' SI propensities this year.

Individual time functions $h_t\left(x_i\right)$ are included to control for time-varying confounding factors with a geographical and/or individual dimension. They are modeled as large numbers of group-specific time-varying dummy variables, e.g., in the form of separate time-dummies for each travel-to-work area or for each neighborhood in Norway, and/or separate time-dummies for groups defined by combinations of birth-year, gender, and educational attainment. In some cases, they also include directly observed covariates, e.g., in the form of indicators for local labor market fluctuations. Their specific formulation vary across different models (as will be explained below), but they are defined on the basis of persons' *initial* characteristics. In the main part of our analysis, we do not exploit information on e.g., migration or additional educational attainment during our observation period, as we expect that such events to some extent are endogenous responses to changes in labor market status (including transitions to social insurance dependency).

Note that although this setup disentangles endogenous interactions from predetermined contextual effects, e.g., related to within-peer-group-correlations in values, preferences, and abilities, it cannot fully separate endogenous interactions from *time-varying* contextual effects. For example, if for some reason an exogenous change in norms/attitudes occurs within a peer-group, this may result in a corresponding change in the affected individuals' observed SI use. We will then not be able to say with certainty whether the subsequently identified contamination effects represent endogenous interactions (i.e., are caused by the peers' actual use of SI) or contextual interactions (i.e., are caused by the peers' changed norms/attitudes).

To estimate Equation (4), we use a fixed effect "within-estimator" that centers the model along several dimensions, and thus avoids estimating parameters that are not of direct interest to us.[3] As a consequence, the model eventually estimated by OLS contains a residual that incorporates a covariate-adjusted individual mean (over all years), and is thus not completely exogenous with respect to the lagged dependent variable (see, e.g., Cameron and Trivedi (2005, p. 764)). Consistency requires that the average residual is small relative to each period's residual, which again requires that the number time periods is large. To assess the potential bias in our case, we have as part of a series of robustness exercises, also estimated Equation (4) with an instrumental variable (2SLS) technique proposed by Anderson and Hsiao (1981). We then rely on first-differencing (instead of mean-centering) to get rid of the person-fixed effect, and instrument the resultant lagged differences $\left\{(y_{i,t-1}-y_{i,t-2}),(\overline{y}_{g,-i,t-1}-\overline{y}_{g,-i,t-2})\right\}$ with their second lag levels $\left\{y_{i,t-2},\overline{y}_{g,-i,t-2}\right\}$.[4] As we show below, it turns out that the first-difference 2SLS estimates of peer effects are somewhat larger, than the fixed effects OLS estimates, though the differences are not statistically significant.

To examine peers' influence on the choice of particular SI program, we also estimate models where we distinguish between the disability-related and the unemployment-related programs (see Section IV). Let $y_{i,t}^{P}$ be the number of months individual *i* claimed benefits of type *P* (=*H*(ealth),*U*(nemployment)) in year *t*. The statistical models then take the form:

---

[3] Due to the large number of observations (up to around 16 million person-years, see next section) and the large number of dummy variables (around 623,000 in the most flexible specification) in addition to the person-fixed effects, estimation raises some computational challenges. We have used a novel algorithm based on The Method of Alternating Projections as described in Gaure (2013) and implemented in the R-package "lfe"; see http://cran.r-project.org/web/packages/lfe/citation.html.

[4] The reason why we also instrument the lagged differenced peer variables is that if there is a same-year peer effect in the true DGP, the differenced residual $(u_{i,t}-u_{i,t-1})$ will be correlated with $(\overline{y}_{g,-i,t-1}-\overline{y}_{g,-i,t-2})$.

$$y_{i,t}^{P}=\alpha_{i}^{P}+\rho_{U}^{P}y_{i,t-1}^{U}+\rho_{H}^{P}y_{i,t-1}^{H}+\lambda^{P}h_{t}\left(x_{i}\right)+\sum_{g\in G}(\theta_{Ug}^{P}\overline{y}_{g,-i,t-1}^{U}+\theta_{Hg}^{P}\overline{y}_{g,-i,t-1}^{H})+u_{it}^{P},\ P=H,U \quad (6)$$

where $(\overline{y}_{g,-i,t}^{U},\overline{y}_{g,-i,t}^{H})$ are the averages (excluding individual *i*) in peer group *g*.

In the next subsections, we first examine interaction effects within three different types of networks separately; i.e., neighborhoods, schoolmates, and ethnic minorities. We then present a brief assessment of some underlying mechanisms, based on separate analyses of SI entry and continuation decisions. In principle, we could have examined all types of networks simultaneously. However, as we explain below, the analysis of each network type requires different cuts and adaptations of the data and the models.

A. Neighbors

We start out examining the impacts of social insurance dependency within residential areas. The purpose is to examine the degree to which SI claim propensities spread endogenously within local communities and to which extent such interaction effects depend on geographical and relational distance. The latter is measured by differences in age, gender, and educational attainment. To avoid endogenous geographical sorting, our analysis is based on recorded address at the start of our analysis period; i.e., in 1992. To reduce the potential attenuation bias caused by subsequent out-migration, we limit the analysis in this subsection to persons belonging to the 1942-1960 birth cohorts, implying that they were between 32 and 50 years old – and hence reasonably settled – at the time of peer group construction in 1992.[5] We also limit the analysis to persons born in Norway, to avoid overlap with a separate analysis of the immigrant population in a later subsection.

We examine peer effects at three geographical levels; neighborhoods, local areas, and municipalities. Our definition of neighborhoods correspond to the so-called "basic statistical

[5] In our data, 58 % of the individuals lived in exactly the same neighborhood in 2008 as they did in 1992.

units" ("grunnkretser") used by Statistics Norway. They are designed to resemble genuine neighborhoods, and contain residences that are homogeneous with respect to location and type of housing.[6] There are 13,700 basic statistical units in Norway, each populated by around 350 individuals on average. Each neighborhood is part of a somewhat larger "local area". Given the geographical proximity, we would expect there to be some room for social interaction between residents of neighboring neighborhoods, although not to the same extent as for same-neighborhood residents. The local areas correspond to the so-called "statistical tracts" ("delområder"), drawn up by Statistics Norway. They are designed to encompass neighborhoods that naturally interact, e.g., by sharing common service/shopping center facilities. A typical local area comprises around 8-9 neighborhoods and 3,100 inhabitants. Local areas are again part of municipalities. There is likely to be some interaction between people living in different local areas in the same municipality also, but less than between people living in the same neighborhood or local area. A typical municipality consists of 3-4 local areas and 11,700 inhabitants.

It follows that we would expect genuine peer effects to be stronger within neighborhoods than within local areas, and stronger within local areas than within municipalities. To ensure that the peer groups in local areas and municipalities are directly comparable to those in the neighborhood, in terms of size as well as composition, we construct them artificially by conducting a one-to-one exact-match sampling; i.e., for each person in *i's* own neighborhood, we draw one person from the local area (outside own neighborhood) and one from the municipality (outside own local area), respectively, who is of the same gender, has the same age (+/- one year), and has exactly the same education.[7] Finally, as part of a placebo analysis, we also

[6] For a more thorough description of the neighborhood concept and other geographical entities used in this paper, see Statistics Norway (1999).

[7] If we find more than one match satisfying these criteria, we draw one of them randomly. If we do not find matches at all geographical levels, the person in question is dropped from the peer group (7.5 % of individuals).

assemble a matched group of "peers" from a different part of the country (defined as being from a non-neighboring county).

In total, there are around 1 million individuals included in this part of the analysis, each of them contributing 16 annual observations (the 1992-observations are lost due the inclusion of the lagged variables); see Table 1. This leaves us with a total number of more than 16 million person-year observations. On average, the persons in our dataset claim social insurance benefits in around 2.7 months each year.

Table 1. Descriptive statistics – Neighborhoods (1942-1960 cohorts)

Number of individuals	1,002,705
Number of neighborhoods	11,828
Average size of the neighborhood (based on observations for individuals)	343.9
Mean annual number of months with SI benefits of any kind	2.68
Mean annual number of months with disability-related benefits	2.37
Mean annual number of months with unemployment-related benefits	0.37
Individuals with 0 benefit months all years (%)	18.9
Individuals with 12 benefit months all years (%)	5.2

Note: The sum of months with disability- and unemployment-related benefits may exceed the total number of months, since it is possible to claim both type of benefits in the same month.

In a baseline model, the vector of time-varying control variables $h_t(x_i)$ includes separate year-dummies for each travel-to-work area (TWA) in Norway and separate year-dummies for each combination of birth-year, sex, and education (the latter with 15 categories reflecting both the level and the type of education).[8] There are 90 TWAs in Norway, defined by Statistics Norway to ensure that persons living in each of these areas operate in a common labor market and have, thus, been subject to the same geographical fluctuations in labor market tightness over time. However, to account for the possibility of labor market fluctuations operating at even lower geographical levels, $h_t(x_i)$ also includes indicators for neighborhood-specific shocks. More specifically, we include an annual downsizing indicator, which is equal

[8] With this specification, we can obviously not distinguish age from time effects, since age and time is perfectly correlated at the individual level; see Biørn *et al.* (2013).

to one if at least two persons belonging to the same neighborhood and working in the same firm register as unemployed in the same year. To account for more general neighborhood-specific economic fluctuations, we first compute nation-wide industry-specific annual transition rates from employment to unemployment for all Norwegian employees.[9] We then use the initial (1992) employment structure in each neighborhood to compute neighborhood-specific weights. Finally, we use these weights, multiplied with the nation-wide time-varying industry specific unemployment risks to compute a variable representing the annual unemployment risks for each neighborhood.

Even with this flexible model, we still cannot rule out the occurrence of confounding shocks – in the form of unaccounted for labor market fluctuations or in the form of changes in the local SI admittance practices. In robustness exercises, we expand the model to comprise separate year dummies for each of the around 450 social insurance districts, for each of the 1,700 local areas, or for each of the 4,700 family-physician practices in Norway, respectively (instead of the 90 TWAs).[10] We also run an additional "placebo" analysis, using annual earnings for non-SI-claimants as the outcome of interest.

[9] We use 12 different industries, based on ISIC codes: i) Farming and fishing, ii) Oil, gas and mining, iii) Manufacturing, iv) Electricity and water supply, v) Construction, vi) Wholesale and retail trade, hotels and restaurants, vii) Transport, storage and communication, viii) Finance, insurance and real estate, ix) Public administration and defense, x) Schools and education, xi) Health services, and xii) Other.

[10] Separate year-dummies for each family-physician group are included to control for possible confounding factors related to changes in the local physicians' lenience/strictness with respect to certifying disability-related SI claims. Registers with information of physician-patient-linkages are not available before 2001; hence we use the 2001 patient lists in this particular exercise. Social insurance districts follow municipality borders except in the largest cities, where there are multiple social insurance districts. Persons living in the same neighborhood or local area also belong to the same social insurance district.

Table 2. Main estimation results – neighborhoods (standard errors in parentheses)

	I **Total use of SI**	II **Disability-related SI**	III **Unemployment-related SI**
Total use of SI last year			
Own claims (t-1)	0.588*** (0.001)		
Avg. claims among peers			
Neighborhood	0.027*** (0.002)		
Local area (matched)	0.010*** (0.002)		
Municipality (matched)	0.008*** (0.002)		
Rest of country (matched)	0.002 (0.002)		
Disability-related SI last year			
Own claims		0.634*** (0.001)	-0.012*** (0.000)
Avg. claims among peers			
Neighborhood		0.037*** (0.002)	-0.011*** (0.001)
Local area (matched)		0.009*** (0.002)	0.000 (0.001)
Municipality (matched)		0.009* (0.002)	0.000 (0.001)
Rest of country (matched)		0.003 (0.002)	-0.002 (0.001)
Unemployment-related SI last year			
Own claims		-0.000 (0.000)	0.480*** (0.001)
Avg. claims among peers			
Neighborhood		-0.019*** (0.003)	0.044*** (0.003)
Local area (matched)		0.002 (0.002)	0.011*** (0.002)
Municipality (matched)		0.000 (0.003)	0.010*** (0.002)
Rest of country (matched)		-0.004 (0.003)	0.000 (0.002)
No. of time-varying dummy variables	9,892	9,892	9,892
By geography (TWAs)	1,336	1,336	1,336
By individual characteristics (sex, birth-year, education)	8,556	8,556	8,556
R-squared	0.797	0.823	0.534
Adj. R-squared	0.783	0.811	0.503
N (persons)	1,002,705	1,002,705	1,002,705
N (person-year observations)	16,043,280	16,043,280	16,043,280

Note: Individual fixed effects are included in all models. The reported R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1)% level.

Estimation results from a set of baseline models are presented in Table 2. Looking first at the total use of SI (regardless of type) in Column I, we note that there is a significant peer effect associated with neighborhoods estimated to 0.027. With an autoregressive coefficient equal to 0.588, this implies a cumulative peer effect (as computed in Equation (5)) around 17%. Moving on to the neighboring neighborhoods in the local area – while maintaining the size as well as the gender-age-education-composition of the peer group – the size of the effect is cut by approximately two thirds; and moving even farther away within the municipality, it declines further. Hence, we identify a clear pattern of declining peer effects as the geographical distance increases. Looking at a matched group of artificial peers in another part of the country, the “effect” is approximately equal to zero. As an additional “falsification test”, we have also estimated a model where we include the behavior of the matched peer group in another part of the country as the *only* peer variable – i.e., as a substitute for the true peer groups. We then obtained a similar insignificant estimate of -0.001 (standard error 0.002), suggesting that the estimated peer effects in Table 2 do not result from unobserved shocks correlated to the gender-age-education-composition of neighborhoods.

Turning to the separate models for disability-related (Column II) and unemployment-related SI claims (Column III), we find that the direct (same SI type) peer effects are of similar size or larger than the total peer effects, while there are small, but significant, negative “cross effects”. The latter indicates a considerable scope for substitution between the two types of SI, and that peer behavior affects both the overall propensity to claim SI benefits and the type of benefits actually claimed.

As noted above, the identification of peer effects in this paper rests on the assumption that controlled for time-varying covariates any remaining shocks in SI claims do not have a spatial pattern that coincides with our peer group definitions. While we have argued that it is hard to envisage such confounding shocks, we now examine the validity of the assumption more formally through a number of robustness analyses. In this exercise, we focus exclusively on the neighborhood peer effect on total SI use. Our primary strategy is to examine what happens with the estimated peer effect as we include ever more flexibility in the time-varying control functions – in the form of shocks at lower geographical levels or in the form of more differentiated geographical shocks. To check the potential bias arising from the correlation between the lagged dependent variable and the residual, we also report estimates from a first-differenced IV (2SLS) model, where the lagged differences are instrumented by their second lag levels.

Table 3. Robustness. Neighborhood peer effect on total SI use (standard errors in parentheses)

	I	II	III	IV	V	VI
Own claims (t-1)	0.588*** (0.001)	0.588*** (0.001)	0.588*** (0.001)	0.587*** (0.001)	0.588*** (0.001)	0.559*** (0.001)
Avg. claims among neighbors in own neighborhood	0.027*** (0.002)	0.021*** (0.002)	0.015*** (0.002)	0.018*** (0.002)	0.026*** (0.002)	0.057** (0.010)
Implied cumulative peer effect	0.170	0.130	0.092	0.110	0.163	0.337
Geographical year dummy variables						
TWA	1,336					1,336
Municipality		6,586				
Local area			23,131			
Physician (in 2001)				57,526		
Individual year dummy variables						
Gender×education×birth-year	8,556	8,556	8,556	8,556		8,550
Interaction of geographical and individual year dummy variables						
Gender×education×birth-year×TWA					622,996	
Estimation method (OLS/2SLS)	OLS	OLS	OLS	OLS	OLS	2SLS
R squared	0.797	0.797	0.798	0.798	0.805	0.795
Adj. R squared	0.783	0.783	0.784	0.783	0.783	0.780
N (persons)	1,002,705	1,002,705	1,002,705	992,002	1,002,705	1,002,705
N (person-year observations)	16,043,280	16,043,280	16,043,280	15,872,032	16,043,280	15,861,750

Note: Individual fixed effects are included in all models. The reported R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) % level.

The results from the robustness analysis are presented in Table 3. As we introduce more flexibility in the controls for local shocks in Columns II-V, the estimated neighborhood peer effects decline somewhat, but remain statistically significant in all specifications. A point to bear in mind here is that the most flexible models entail the risk of "over-controlling", in the sense that the dummy control vectors absorb some genuine peer effects. It is notable that

that the model's overall explanatory power – as measured by R squared – is virtually unchanged as separate year dummies are introduced at ever lower geographical levels. For example, substituting 57,526 family-physician year dummies (Column IV) for the 1,336 TWA year dummies (Column I) raises the unadjusted R-squared from 0.797 to 0.798.

Switching estimation technique, from fixed effects OLS to first-differenced 2SLS, does change the estimated peer effect considerably, yielding a cumulative effect as high as 0.34; see Column 6. The standard errors also become much larger, however, suggesting that the change may be due to statistical uncertainty. Yet, if anything, the 2SLS-results indicate that our within-estimators may underestimate the neighborhood effect, rather than overestimating it.

While the robustness exercises reported in Columns II-V arguably control fully for conceivable confounders related to, e.g., the gatekeeping practices exercised by local social insurance offices and panel doctors, one may perhaps worry that there could still be some neighborhood-specific labor market developments that are not completely accounted for by the time-dummies defined at higher (or different) geographical levels. One way to check this is to examine whether our neighborhood peer variable is correlated with alternative measures of labor market success within the neighborhood – measures that are insulated from influence of social insurance behavior. Hence, we perform a "placebo" analysis where we use annual labor earnings as the outcome of interest for the 184,159 persons in the dataset who *never* claimed SI during our observation period. This is obviously a highly selected group of persons. However, if there are any remaining unaccounted for time-varying confounders related to economic fluctuations at the neighborhood level, these confounders most likely would affect earnings levels as well as employment levels. Our placebo analysis is based on an individual fixed effects model where we include exactly the same neighborhood SI peer variable as in the regressions above, and also include year dummies by TWA and by individual characteristics (as in Table 3, Column I). As it turns out, we find no effect of the SI peer variable on annual earnings (t-value=0.027; not reported in tables).

Further insights to the nature of the neighborhood peer effects identified for SI claims may be gained by assessing the importance of "relational closeness". If persons interact more with neighbors that are similar to themselves, we may hypothesize that persons are more strongly influenced by persons of same sex and similar age and education than by more dissimilar neighbors. To examine the empirical relevance of this hypothesis, we have re-estimated the baseline model for total SI claims, using a multiple of group-specific averages within own neighborhoods as explanatory variables. To ascertain direct comparability, we weight each group mean by its size relative to the whole neighborhood, such that each coefficient is directly comparable to the overall neighborhood effect reported in Table 2, Column I. The results are presented in Table 4. They confirm that relational closeness is important. Persons respond more strongly to the behavior of similar than dissimilar neighbors, particularly along the dimensions of sex and age; see Columns I-III. Similarity in education, on the other hand, does not appear to be critical for the degree of social interaction among neighbors. As a sort of robustness exercise, Column IV report results for a model where we only include neighbors of the same sex *and* the same age group in the computation of the peer variable (again weighted relative to the size of the whole neighborhood to ensure direct comparability) while including a full set of 177,406 neighborhood×year dummy variables. In this model, the general neighborhood effects are fully absorbed by the neighborhood×year dummies, whereas the estimated peer effect is interpreted as the "extra" effect that would have arisen if all neighbors belonged to the same sex and age group. The result is in line with what we would expect on the basis of group specific estimations reported in Column I-III, and confirms the social interaction interpretation of our findings.

Table 4. Neighborhood peer effect on total SI use. By "relational closeness" (standard errors in parentheses)

	I	II	III	IV
By sex:				
Own claims (t-1)	0.588*** (0.001)			
Average claims among peers:				
Own sex	0.0378*** (0.003)			
Opposite sex	0.0265*** (0.003)			
By age:				
Own claims (t-1)		0.588*** (0.001)		
Average claims among peers:				
Younger		0.004* (0.004)		
Same age (±5 years)		0.046*** (0.004)		
Older		0.026*** (0.005)		
By education:				
Own claims (t-1)			0.588*** (0.001)	
Average claims among peers:				
Lower education			0.031*** (0.004)	
Similar education (see note below)			0.033*** (0.004)	
Higher education			0.032*** (0.005)	
Same sex and age group:				
Own claims (t-1)				0.587*** (0.001)
Average claims among same sex and same age neighbors				0.038*** (0.006)
No. of time-varying dummy variables	9,902	9,902	9,902	185,972
By geography (TWAs)	1,336	1,336	1,336	
By geography (neighborhoods)				177,406
By individual characteristics (sex, birth-year, education)	8,566	8,566	8,566	8,566
R-squared	0.800	0.800	0.800	0.800
Adj. R-squared	0.786	0.786	0.786	0.784
N (persons)	1,002,705	1,002,705	1,002,705	1,002,705
N (person-year observations)	16,043,280	16,043,280	16,043,280	16,043,280

Note: Individual fixed effects are included in all models. The reported R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Comparison of education levels is based on three groups: i) Less than 11 years (primary education only), ii) 11-13 years (lower or upper secondary), iii) more than 13 years (college, university). Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

Another way of addressing the importance of relational closeness is to estimate peer effects separately for neighborhoods that are different with respect to the general level of social interaction among neighbors. This is of course not observed in administrative register data. We may assume, however, that social interaction is more frequent in neighborhoods with, say, many native families with children and many married (settled) couples, than in neighborhoods with many singles, many students, and a large fraction of immigrants. Based on this idea, we thus compute neighborhood-specific social interaction indicators, which we subsequently use to classify neighborhoods in terms of expected interaction levels.[11] Finally, we estimate our baseline model separately for neighborhoods with particularly low and particularly high expected interaction levels (the 25 % most extreme neighborhoods at each tail of the distribution). What we then find is that the estimated own neighborhood peer effect is 0.028 (standard error 0.004) in neighborhood with high expected social interaction and 0.021 (standard error 0.005) in neighborhood with low expected interaction. Hence, according to these estimates, the peer effect is approximately 25 percent larger in high-interaction neighborhoods.[12]

B. Schoolmates

We now turn our attention to networks consisting of persons who went to the same junior high school at the same point in time. Junior high school in Norway is a three-year track, normally attended at age 13-15. The total group of school mates during this period thus consists of five birth-cohorts; those at the same age, and those born up to two years before and two years after. We start out this subsection examining the peer effects present within this complete group. We then take a closer look at the importance of relational closeness, in this case measured by differences in class-levels (age) and gender. Due to data limitations, we can only use a subset of our analysis population in this part of the analysis, namely those born between 1961 and 1971 (11 cohorts). To ensure that different birth-cohorts really went to dif-

[11] The classification is based on all residents in the neighborhoods in 1992, also those who are not included in the analysis in this section. For each resident, we compute a variable which is equal to 1 if the person has at least one child below 19 years *or* is above 45 years *and* married *and* is not a student *or* an immigrant. Our social interaction score is then the average of this variable for all residents in the neighborhood.

[12] As a sort of plausibility-test, we have also estimated separate models for those who lived in the same neighborhood throughout our data period (58 %) and those who did not; i.e., we have performed the whole (baseline) analysis on these two populations separately. The estimated peer effect is then more than twice as large for the "stayers" (point estimate 0.024, standard error 0.002) than for the "movers" (point estimate 0.011, standard error 0.003). Note that both estimates are smaller than the one estimated in the baseline model, reflecting that the peer groups are much smaller in this analysis.

ferent classes, we also require the group of “levelmates” to comprise at least 30 persons. Finally, we remove siblings from each person’s peer group. In total, we construct data for 5,850 schoolmate groups. Descriptive statistics are provided in Table 5.

Note that common shocks related to the schooling experience – such as being subject to a particularly good (or bad) principal or teacher – will not represent a confounder in this analysis, since such events took place several years before our outcome period and, hence, presumably would be captured by the individual-fixed effects. It may still be the case, though, that persons who went to the same class/school are affected by the same shocks later on, as many of them continue to reside in the geographical area they grew up in. We control for this potential confounding factor in the same way as in the preceding subsection; i.e., by including separate year dummy variables for each travel-to-work area (TWA) based on the address recorded at the start of the outcome period. In robustness analyses, we introduce year-dummies at lower geographical levels, all the way down to the neighborhood.

Table 5. Descriptive statistics – Schoolmates (1961-1971 cohorts)

Number of individuals	515,666
Number of schoolmate groups	5,850
Average number of schoolmates (taken over individuals included in the data)	524
Mean annual number of months with SI benefits of any kind	1.81
Mean annual number of months with disability-related benefits	1.31
Mean annual number of months with unemployment-related benefits	0.60
Individuals with 0 benefit months all years (%)	17.1
Individuals with 12 benefit months all years (%)	1.0

The results indicate significant peer effects among former schoolmates: see Table 6. Looking first at the total use of SI in Column I, the peer effect is estimated to 0.059, which together with the autoregressive parameter implies a cumulative effect of 25 %. Moving on to the separate models for disability-related (Column II) and unemployment-related claims (Column III), we again find a pattern of positive direct effects and negative cross effects.

Table 6. Main estimation results – schoolmates (standard errors in parentheses)

	I **Total use of SI**	II **Disability-related SI**	III **Unemployment-related SI**
Total use of SI last year			
Own claims	0.483*** (0.001)		
Avg. claims among former schoolmates	0.059*** (0.007)		
Disability-related SI last year			
Own claims		0.584*** (0.001)	-0.013*** (0.000)
Avg. claims among former schoolmates		0.043*** (0.008)	-0.013** (0.006)
Unemployment-related SI last year			
Own claims		-0.001 (0.001)	0.400*** (0.001)
Avg. claims among former schoolmates		-0.009 (0.006)	0.067*** (0.007)
No. of time-varying dummy variables	6,887	6,887	6,887
By geography (TWAs)	1,336	1,336	1,336
By individual characteristics (sex, birth-year, education)	5,551	5,551	5,551
R-squared	0.663	0.723	0.489
Adj. R-squared	0.640	0.704	0.454
N (persons)	515,666	515,666	515,666
N (person-year observations)	8,250,656	8,250,656	8,250,656

Note: Individual fixed effects are included in all models. The reported R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

Robustness is evaluated in Table 7, where we control for potential time-varying local confounders at lower geographical levels. The estimated peer effect decline slightly when we use separate year dummies at local area or the neighborhood levels. It is notable, though, that the estimated effects are unchanged when we substitute more than 185,000 neighborhood-year-fixed effects for 23,000 locale-area-year-fixed effects. The estimates again rise a bit when we use the first-differenced 2SLS estimator rather than the fixed effects OLS.

Table 7. Robustness. Schoolmate peer effect on total SI use (standard errors in parentheses)

	I	II	III	IV	V
Own claims (t-1)	0.483*** (0.001)	0.483*** (0.001)	0.483*** (0.001)	0.483*** (0.001)	0.468*** (0.002)
Avg. claims among former schoolmates	0.059*** (0.007)	0.050*** (0.007)	0.050*** (0.007)	0.046*** (0.008)	0.067*** (0.017)
Implied cumulative peer effect	0.249	0.207	0.207	0.189	0.271
Geographical year dummy variables					
TWA	1,336				1,336
Local area		23,296			
Neighborhood			184,621		
Individual year dummy variables					
Gender×Education×birth-year	5,551	5,551	5,551		5,550
Interaction of geographical and individual year dummy variables					
Gender×Education×birth-year×TWA				375,121	
Estimation method (OLS/2SLS)	OLS	OLS	OLS	OLS	2SLS
R squared	0.663	0.664	0.671	0.678	-
Adj. R squared	0.640	0.640	0.640	0.639	-
N (persons)	515,666	515,666	515,666	515,666	515,666
N (person-year observations)	8,250,656	8,250,656	8,250,656	8,250,656	7,734,990

Note: Note: Individual fixed effects are included in all models. R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

Most adults probably have little (if any) contact with the majority of the persons they went to junior high-school with. Hence, by including all former schoolmates in our peer measure, we clearly include a large number of irrelevant persons. It may therefore be of some interest to distinguish “close” from more “distant” peers. In particular, we would guess that former classmates are more likely to maintain a relationship with each other than persons who went to different classes or levels. And it is also probable that same-sex persons have maintained more contact than persons of different sexes. We examine the issue of relational closeness by estimating separate peer effects based on a same-level-same-sex distinction; see Table 8. Note that we have weighted each group’s SI average with its size relative to the total number of schoolmates (all five cohorts), such that the coefficients are directly comparable to each other and to the total schoolmate effect reported in Table 7. Again, the results indicate that relational closeness is a key factor in understanding social interaction effects. As shown in Column I, the impact of same-level-same-sex peers is much larger than the impact of other schoolmates. And for schoolmates of the opposite sex, we find no significant peer effects at all. As an additional robustness exercise, we can take advantage of the differentiation between close and distant peers to check for possible confounding school-specific developments. In Column II, we report the estimated same-level-same-sex peer effect in a model where we also include separate year dummy variables for each school included in the dataset. While these dummies may absorb some genuine peer effects related to the overall mass of schoolmates, we note that the estimated effect of the presumed closest peers declines only slightly.

Table 8. Total use of SI. Estimated peer effects by relational closeness. Schoolmates. (standard errors in parentheses)

	I	II
Own claims	0.483*** (0.001)	0.483*** (0.001)
Avg. claims among former schoolmates		
Same level same sex	0.125*** (0.025)	0.093*** (0.026)
Same level opposite sex	0.023 (0.025)	
1-2 levels above/below same sex	0.053*** (0.012)	
1-2 levels above/below opposite sex	0.012 (0.011)	
No. of time-varying dummy variables	6,887	15,449
By geography (TWAs)	1,336	
By school		9,898
By individual characteristics (sex, birth-year, education)	5,551	5,551
R-squared	0.663	0.663
Adj. R-squared	0.640	0.640
N (persons)	515,666	515,666
N (person-year observations)	8,250,656	8,250,656

Note: Note: Individual fixed effects are included. R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

C. Ethnic minorities

Some of the most influential existing studies on social insurance interaction effects are based on data for ethnic minorities (Bertrand *et al.*, 2000; Aizer and Currie, 2004; Åslund and Fredriksson, 2009). We follow up on this literature by looking at SI use among immigrants from low-income countries.[13] Our focus is on immigrants who reside in areas where there are sufficient numbers of other immigrants from the same country for a network of some size to be established. More specifically, we define an ethnic minority network as a group of immigrants

[13] We disregard immigrants from high-income countries here, both because they do not tend to be concentrated in particular geographical areas and because they do not tend to reside permanently in Norway.

from the same origin country who resided in the same local area in 1992 (the "neighborhoods" discussed above are too small for this purpose). To be included in the analysis, we require a network size of minimum 10 persons. Based on this strategy, we end up with 23,306 persons, divided between 746 local immigrant networks; see Table 9 for descriptive statistics.

Table 9. Descriptive statistics – Ethnic minorities (1942-1974 cohorts)

Number of individuals	23,306
Number of immigrant networks	746
Average size of immigrant network (taken over individuals in included in the data)	101.3
Mean annual number of months with SI benefits of any kind	3.87
Mean annual number of months with disability-related benefits	2.52
Mean annual number of months with unemployment-related benefits	1.53
Individuals with 0 benefit months all years (%)	9.3
Individuals with 12 benefit months all years (%)	3.4

One could imagine that the social interaction effects decrease with geographical distance for immigrants as well as for natives, suggesting that we should examine how the estimated effects change as we substitute close groups with more distant ones (but with the same nationality). Our data impose some limitations, however, as nationality networks of the required size are typically located closely together. Instead, we use immigrants from other low-income countries as candidates for more "distant" peers. In addition, we look at how immigrants are affected by SI use among natives within the same local area. Again, we compose the groups of other immigrants and natives such that they are of equal size and have similar characteristics as the person's own same-nationality network. We are not able to obtain exact matches of the same quality as those used in the neighborhood analysis above, and the relatively low number of observations available for this analysis also implies that we cannot "afford" to drop observations with imperfect matches. Hence, while we have a perfect matching on sex, we allow for poorer matches on age and educational attainment. We are also not able to control for time-varying confounders at a lower level than travel-to-work areas. Note, however, that immigrants from different low-income countries typically work in similar sectors of the economy, with a domination of low-skill service sector jobs (Bratsberg *et al*., 2010); hence if uncontrolled-for confounding factors remain at the local level, they would presumably affect persons from different low-income countries in a similar fashion.

Table 10. Main estimation results – nationalities (standard errors in parentheses)

	Total use of SI		Disability-related SI	Unemployment-related SI
	I	II	III	IV
Total use of SI last year				
Own claims	0.535*** (0.002)	0.557*** (0.006)		
Avg. claims among peers				
Immigrants from same source country	0.070*** (0.009)	0.081*** (0.021)		
Immigrants from other low-income country (matched)	-0.012 (0.007)	-0.014 (0.018)		
Natives (matched)	-0.008** (0.012)	-0.002 (0.031)		
Disability-related SI last year				
Own claims			0.612*** (0.003)	-0.008*** (0.002)
Avg. claims among peers				
Immigrants from same source country			0.051*** (0.009)	-0.027*** (0.010)
Immigrants from other low-income country (matched)			-0.014** (0.007)	-0.001 (0.008)
Natives (matched)			-0.007 (0.011)	0.003 (0.012)
Unemployment-related SI last year				
Own claims			0.003* (0.002)	0.479*** (0.003)
Avg. claims among peers				
Immigrants from same source country			-0.023*** (0.007)	0.117*** (0.010)
Immigrants from other low-income country (matched)			-0.008 (0.007)	-0.013 (0.008)
Natives (matched)			-0.002 (0.013)	-0.026* (0.015)
Estimation method (OLS/2SLS)	OLS	2SLS	OLS	OLS
No. of time-varying dummy variables	18,122	18,121	18,122	18,122
By geography (TWAs)	931	931	931	931
By individual characteristics (sex, birth-year, education)	17,191	17190	17,191	17,191
R-squared	0.719		0.603	0.603
Adj. R-squared	0.684		0.554	0.554
N (persons)	23,306	23,306	23,306	23,306
N (person-year observations)	372,896	349,590	372,896	372,896

Note: Note: Individual fixed effects are included in all models. R-squared is a goodness-of-fit measure for the complete model, including the individual-fixed effects. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

Table 10 presents the results, also including the 2SLS results for the main model. According to the OLS estimates, there is a significant peer effect among immigrants from a common source country – stronger than what we have found to be the case for neighbors in general and former schoolmates. The cumulative total SI peer effect is around 38%. On the other hand, we find no peer effects among immigrants from different source countries, and indications of a small *negative* effect of natives' SI claims. The 2SLS point estimates are almost identical to OLS for these networks, yet with much larger standard errors. Turning to the separate estimations for disability-related and unemployment related claims, we again find patterns of large positive direct effects and negative cross-effects.

D. Mechanisms

Peer effects can be driven by information-sharing and by propagation of norms/stigma. While we would expect information-sharing to be relevant for entry decisions only (or at least primarily), more general norm-effects are relevant for both entry and continuation decisions. Hence, by examining peers' influence on entry and continuation separately we may gain some understanding of the underlying mechanisms. To do this, we have dichotomized the outcome variable used in the previous three subsections (=1 for positive SI claims during a year, =0 otherwise), and split our dataset into three (partly overlapping) parts. To examine inflow into disability-related and unemployment-related SI, we use annual observations for which there were no SI claims last year $(y_{i,t-1} = 0)$. To examine continuation of disability-related SI claims, we use observations for which there were some disability-related SI claims last year. And to examine continuation of unemployment-related SI claims, we use observations for which there were some unemployment-related claims last year (persons with both disability- and unemployment-related claims last year are included in both the two latter groups). We then redo the main statistical analyses by program type – for neighbors, schoolmates, and ethnic minorities. Except that the lagged dependent variable drop out of the analyses, the statistical models, the peer variables, and the control variables are exactly the same as in previous subsections (conf. Tables 2, 6, and 10, respectively); i.e., OLS (in this case linear probability models) with person-fixed effects and individual time-controls (based on TWA and combinations of gender, age, and education).

Table 11. Peer effects on inflow to and continuation of SI benefit claims (standard errors in parentheses)

	Disability-related SI		Unemployment-related SI	
	I Inflow	II Continuation	III Inflow	IV Continuation
Avg. claims among neighbors				
Disability-related	0.0043*** (0.0004)	0.0005 (0.0007)	-0.0004** (0.0002)	0.0017 (0.0021)
Unemployment-related	-0.0018*** (0.0006)	0.0047*** (0.0011)	0.0018*** (0.0003)	0.0070*** (0.0026)
N (persons)	917,055	759,075	917,055	296,009
N (person-year observations)	10,668,525	4,613,879	10,668,525	1,151,994
Avg. claims among school-mates				
Disability-related	0.0055*** (0.0013)	0.0077** (0.0033)	-0.0008 (0.0008)	0.0007 (0.0060)
Unemployment-related	-0.0008 (0.0013)	0.0016 (0.0039)	0.0095*** (0.0010)	0.0177*** (0.00551)
N (persons)	500,035	374,021	500,035	268,058
N (person-year observations)	8,783,043	1,725,483	5,783,043	1,041,105
Avg. claims among immigrants from same source country				
Disability-related	0.0027 (0.002)	0.0015 (0.0029)	-0.001 (0.002)	0.002 (0.005)
Unemployment-related	-0.006*** (0.002)	-0.0005 (0.003)	0.0097*** (0.001)	0.0085*** (0.0028)
N (persons)	20,704	18,114	20,704	16,186
N (person-year observations)	184,994	117,254	184,994	92,719

Note: Dependent variable is a dummy variable indicating whether or not SI was received at all in a given year. The peer variables are the same as those used in the previous three subsections. Control variables are also the same as those reported in Tables 2, 6, and 10, respectively, with time-varying dummy controls defined at the TWA level and by combinations of gender, age, and education. Individual fixed effects are included in all models. Standard errors are clustered at the peer group level. *(**)(***) indicates significance at the 10(5)(1) percent level.

The estimated coefficients associated with the main peer variables are presented in Table 11. Two interesting patterns emerge. The first is that the direct (same-program) peer effects are positive and in most cases significant for both entry and continuation decisions. This suggests that the peer-effects are not only driven by information-sharing; even "experienced" claimants respond to peer behavior. The second is that the negative cross-program peer effects are entirely driven by entry-decisions. For the continuation decisions, the cross-program peer effects are either *positive* or zero. This suggests that peers have significant influence on a new claimant's "choice" of program by sharing experiences regarding entitlement and application procedures. But once a person has already become a claimant, the peer influences are dominated by more general work-morale effects; higher SI claims among peers increases the payoff associated with own continuation even when the increase stems from a different program type.

VI. Conclusion

We have shown that there are significant social interaction effects in the use of social insurance (SI) benefits in Norway. Exogenous changes in SI dependency tend to be enlarged by self-enforcing group-behavior, implying the existence of a social multiplier. To avoid the problems of endogenous group formation, our analysis has been based on pre-determined peer groups – within which the majority of the members presumably have little or no contact with each other. We nevertheless estimate cumulative knock-on effects associated with exogenous changes in SI claims within these networks amounting to at least 10-15%. Our estimates may be interpreted as lower bounds on peer effects prevailing in more closely knitted networks of genuine friends and actually interacting neighbors.

An important finding of our paper is that peer behavior not only affects individuals' overall propensity to claim social insurance benefits, but also the *type* of program to which claims are directed. For example, a rise in a peer group's disability insurance claims increases the group-members propensity to claim disability insurance, but at the same time it significantly reduces their propensity to claim unemployment benefits or social assistance (and vice versa). The negative "cross-effect" only applies to initial non-claimants, though, suggesting that the information value of experience-sharing within networks is empirically important.

Previous empirical evidence has shown that there is indeed a significant overlap in the caseloads of different social insurance programs in Norway, and that job loss more than doubles the risk of entry into disability-related SI programs (Bratsberg *et al.*, 2013). Our own findings provide further evidence on the substitutability between SI programs, and indicate that that the dividing line drawn between them is path dependent. They also indicate that previous empirical findings reported for peer effects within specific programs (see Section II) must be interpreted with some care, as they may reflect a combination of contamination in the overall use of SI and a substitution for other SI programs.

The methodological approach used in this paper has been designed to identify and estimate local social propagation mechanisms based on the timing – rather than the occurrence – of claims; and we have argued that we have done so in a way that convincingly and robustly distinguishes social interactions from other sources of within-group correlations. We have identified a conspicuous tendency for estimated interaction effects to rise with measures of relational closeness in a way that, given our vector of control variables, is unlikely to have been caused by confounding shocks. Any social contagion operating at the aggregate or regional level, however, for example through an effect of overall SI propensity on the disutility/stigma of claiming SI benefits, have been effectively "controlled away" by the use of separate year dummy variables for different travel-to-work areas. We have done this *not* because we believe that such aggregate/regional effects are empirically irrelevant, but because we see no way to convincingly disentangle them from other sources of time changes in SI dependence rates. Indeed, we will argue that the identification of social multipliers at local levels may be indicative of such effects being present at the aggregate level as well.

The complementarities in individual behaviors exposed in our empirical analysis can potentially explain why large regional differences in SI claim patterns persist and why we sometimes witness time-trends with no apparent observed cause. In particular, they may shed some new light on the conspicuous (but largely unexplained) rise in Norwegian disability insurance claims between 1995 and 2004, which apparently coincided (perhaps with a small time lag) with a steep decline in unemployment insurance claims (see Figure 2, panels (c) and (e) in Section IV above). The fall in unemployment was no doubt triggered by cyclical factors, although the results presented in this paper indicate that it may have gained momentum through social interaction effects. More importantly in the present context, it may have contributed to the subsequent rise in disability-related insurance claims, which in turn also gained momentum through self-enforcing network effects. As unemployment dropped, it became relatively more stigmatizing to claim unemployment insurance benefits, and, as a result, the demand for alternative disability-related benefits rose.

The empirical strategy used in this paper cannot perfectly distinguish endogenous from contextual social interactions. However, since we have used a combination of individual fixed effects and stable predetermined peer groups, we can rule out that the estimated peer effects reflect contextual interactions arising from predetermined social factors. To the extent that effects represent endogenous social interactions, there are important policy implications with respect to the cost-benefit-assessment of strategies affecting the SI caseloads. If governments can find ways to reduce the social insurance rolls directly – e.g., by tightening gatekeeping, increasing rehabilitation efforts, reducing benefit levels, or by expanding activation programs – they can expect a significant “bonus” reduction through the social multiplier. This implies that strategies to get individuals off the SI roll may be cost effective even when the direct costs exceed the benefits for each individual claimant. Furthermore, the mere existence of (sizeable) social interaction effects can be interpreted as evidence that moral hazard problems are empirically relevant: SI claims are not triggered by exogenous job loss or health shocks alone; they are the result of individual choices made on the basis of individual preferences. And these preferences apparently incorporate a malleable social norm.

References

Aizer, Anna, and Janet Currie. 2004. “Networks or Neighborhoods? Correlations in the Use of Publically-funded Maternity Care in California.” *Journal of Public Economics* 88(12):2573-85.

Anderson, T. W. and Cheng Hsiao. 1981. “Estimation of Dynamic Models with Error Components.” *Journal of the American Statistical Association* 76: 598-606.

Åslund, Olof and Peter Fredriksson. 2009. “Peer Effects in Welfare Dependence. Quasi-Experimental Evidence.” *Journal of Human Resources* 44(3):798-825.

Autor, David H. and Mark G. Duggan. 2003. “The Rise in the Disability Rolls and the Decline in Unemployment.” *The Quarterly Journal of Economics* 118(1):157-205.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. “Network Effects and Welfare Cultures.” *Quarterly Journal of Economics* 115(3):1019-55.

Biørn, Erik, Simen Gaure, Simen Markussen, and Knut Røed. 2013.”The Rise in Absenteeism: Disentangling the Impacts of Cohort, Age and Time.” *Journal of Population Economics* 26(4):1585-1608.

Black, Dan, Kermit Daniel, and Seth Sanders. 2002. “The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust.” *American Economic Review* 92:27–50.

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides. 2010. “*Identification of Social Interactions*.” In Benhabib, J., Bisin, A. and Jackson, M (eds.): Handbook of Social Economics. Elsevier, North-Holland.

Bradley, Steve, Colin P. Green, Gareth Davis Leeves. 2007. “Worker Absence and Shirking: Evidence from Matched Teacher-School Data.” *Labour Economics* 14:319–34.

Bratberg, Espen, Øivind Anti Nilsen, and Kjell Vaage. 2012. “Intergenerational correlation in disability benefits receipt.” Working Paper 07-2012, Dept of Economics, University of Bergen.

Bratsberg, Bernt, Elisabeth Fevang, and Knut Røed. 2013. “Job Loss and Disability Insurance.” *Labour Economics* 24:137–50.

Bratsberg, Bernt, Oddbjørn Raaum, and Knut Røed. 2010. “When Minority Labor Migrants Meet the Welfare State.” *Journal of Labor Economics* 28(3):633-76.

Brock, William A. and Steven N. Durlauf. 2000. “Interaction-Based Model.” In Heckman, J. and Leamer, E. (Eds.) Handbook of Econometrics, Vol. 5, North-Holland.

Burkhauser, Rkichard V. and Mary C. Daly. 2011. “*The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change*.” AEI Press, Washington D.C.

Cameron, A. Colin and Pravin K. Trivedi. 2005. “Microeconometrics – Methods and Applications.” Cambridge University Press.

Conley, Timothy G and Giorgio Topa. 2002. “Socio-Economic Distance and Spatial Patterns in Unemployment.” *Journal of Applied Econometrics* 17:303-27.

Cont, Rama and Matthias Löwe. 2010. “Social Distance, Heterogeneity, and Social Interactions.” *Journal of Mathematical Economics* 46: 572-90.

Dahl, Gordon B., Andreas R. Kostøl, and Magne Mogstad. 2013. “Family Welfare Cultures." *NBER Working Paper* 19237.

Duggan, Mark and Scott A. Imberman. 2006. "Why Are Disability Rolls Skyrocketing?" In D. Cutler and D. Wise (eds.): *Health in Older Ages: The Causes and Consequences of Declining Disability among the Elderly*. University of Chicago Press.

Durlauf, Steven N. 2004. "Neighborhood Effects." In Handerson and Thisse (Eds): *Handbook of Urban and Regional Economics*, Vol. IV, North-Holland, Amsterdam.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller. 2011. "The Demand for Social Insurance: Does Culture Matter?" *Economic Journal* 121:F413-48.

European Commission. 2011. "*Your social security rights in Norway*." The European Commission (http://ec.europa.eu/social-security-directory)

Fevang, Elisabeth, Knut Røed, Lars Westlie, and Tao Zhang. 2004. "Veier inn i, rundt i, og ut av det norske trygde- og sosialhjelpssystemet." Rapport 6/2004, Stiftelsen Frischsenteret for samfunnsøkonomisk forskning.

Gaure, Simen. 2013. "OLS with Multiple High Dimensional Category Variables." *Computational Statistics & Data Analysis* 66:8–18.

Glaeser, Edward L., Bruce I. Sacredote, and Jose A. Scheinkman. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1(2-3)345-53.

Halvorsen, Knut and Steinar Stjernø. 2008. "*Work, Oil and Welfare*." Universitetsforlaget, Oslo.

Hesselius, Patrik, Peter Nilsson, and Per Johansson. 2009. "Sick of Your Colleagues'

Absence?" *Journal of the European Economic Association* 7(2-3):583-94.

Ichino, Andrea and Giovanni Maggi. 2000. "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *Quarterly Journal of Economics* 115(3):1057-90.

Ioannides, Yannis M. and Linda D. Loury. 2004. "Job Information Networks, Neighborhood Interactions and Inequality." *Journal of Economic Literature* 42:1056-93.

Lindbeck, Assar. 1995. "Hazardous Welfare-State Dynamics." *American Economic Review*, Papers and Proceedings 85(2):9-15.

Lindbeck, Assar, Sten Nyberg, and Jörgen Weibull. 1999. "Social Norms and Economic Incentives in the Welfare State." *Quarterly Journal of Economics* 114(1):1-35.

Lindbeck, Assar, Sten Nyberg, and Jörgen Weibull. 2003. "Social Norms and Welfare State Dynamics." *Journal of the European Economic Association* 1:533-42.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60(3):531-42.

Markussen, Simen, Knut Røed, and Ole J. Røgeberg. 2013. "The Changing of the Guards - Can Family Doctors Contain Worker Absenteeism?" *Journal of Health Economics* 32(6):1230–39.

Marsden, Peter V. 1982. "Homogeneity in Confiding Relations." *Social Networks* 10(1):57-76.

McCoy, John L., Miles Davis, and Russel E. Hudson. 1994. Geographic Patterns of Disability in the United States. *Social Security Bulletin* 57(1):25-36.

Moffitt, Robert. 1983. “An Economic Model of Welfare Stigma.” *The American Economic Review* 73(5):1023-35.

OECD. 2010. “Sickness, Disability and Work – Breaking the Barriers. A Synthesis of Findings across OECD Countries.” OECD, Paris.

Rege, Mari, Kjetil Telle, and Mark Votruba. 2009. “The Effect of Plant Downsizing on Disability Pension Utilization.” *Journal of the European Economic Association* 7(5):754–85.

Rege, Mari, Kjetil Telle, and Mark Votruba. 2012. “Social Interaction Effects in Disability Pension Participation – Evidence from Plant Downsizing.” *Scandinavian Journal of Economics* 114(4):1208–39.

Statistics Norway. 1999. ”Regionale inndelinger - En oversikt over standarder i norsk offisiell statistikk.” Norges Offisielle Statistikk, Oslo-Kongsvinger.

Stutzer, Alois and Rafael Lalive. 2004. “The Role of Social Work Norms in Job Search and Subjective Well-Being.” *Journal of the European Economic Association* 2(4):696-719.