

# The Association Between Own Unemployment and Violence Victimization Among Female Youths

Michèle Tertilt\*  
Gerard J. van den Berg<sup>†</sup>

October 9, 2014

**Abstract** We estimate the association between the unemployment status of young women and the probability that they are subject to violence, using Swedish population register data covering the period 1999-2008. These data contain the highest-level classification of diagnoses made by medical experts at every individual in-patient and out-patient visit to medical care units, including every contact with a physician. We distinguish between domestic and non-domestic violence. It turns out that unemployed women are significantly more likely to be victimized than employed women with the same individual characteristics. This is mostly reflected in indicators of non-domestic violence and long-run abuse among unemployed female youths.

---

\*University of Mannheim, CEPR and BREAD

<sup>†</sup>University of Mannheim, IFAU-Uppsala, IZA, ZEW and CEPR.

Keywords: conflict, household, relationships, assaults, abuse, health, crime, ICD codes.

JEL codes: I10, I12, J01, J10, J13, J24.

We are grateful to two anonymous Referees for their very useful comments and to Alexander Paul for his research assistance. We thank participants at the ZEW workshop on youth unemployment (in honor of Wolfgang Franz' 70th birthday) for helpful comments. We thank Statistics Sweden and Socialstyrelsen for the use of the micro data registers and Hans-Martin von Gaudecker and Mårten Palme for their efforts to make these data accessible. Tertilt thanks the ERC for financial support. Van den Berg is Alexander von Humboldt Professor of Econometrics and Empirical Economics.

# 1 Introduction

It is generally acknowledged that youth unemployment is costly for society. This concerns first and foremost the direct economic costs in terms of lost opportunities. In addition, youth unemployment may lead young individuals to pursue a career in crime. This has recently been acknowledged in the literature (see e.g. Fougère, Kramarz and Pouget, 2009). What has been less closely investigated is whether, among youths, individual unemployment is associated with a higher incidence of victimization due to crime. For example, if unemployment leads to crime and unemployed individuals are socially clustered then it is conceivable that young unemployed individuals are more exposed to crime than their employed counterparts. In such a scenario, unemployment creates an incentive to become involved in crime and/or it leads individuals to become more violent. A young individual who becomes unemployed may socialize more with unemployed individuals and less with employed individuals. As a result, the individual may be more at risk of being robbed, abused or assaulted. Alternatively, unemployment and exposure to crime among youths may both depend on the level of education (or on other personal characteristics). A high level of education may increase the likelihood of having a conventional lifestyle and a regular career with permanent employment, and it may also increase the extent to which the individual is sheltered from crime.

In this paper we consider youth exposure to violence. Specifically, we estimate the association between individual unemployment and victimization due to violence, among female youths aged 16 to 25. We distinguish between domestic<sup>1</sup> and non-domestic violence. Being subjected to violence is among the most traumatic events conceivable. In the case of domestic violence, victims are often tied to the perpetrator in a relationship of close economic and emotional dependence. This may enhance the trauma and reduce the incentive to report victimization.

A number of empirical studies have used cross-sectional survey data to study the associations we are interested in (see e.g. Tauchen, Witte and Long, 1991, Lloyd 1997, and Bowlus and Seitz, 2006, and references therein). However, these studies do not focus on youths but instead consider unemployment across all ages. To the extent that existing studies consider domestic (rather than non-domestic) violence, they often focus on the employment status of the perpetrator rather than the victim. Some of the existing studies use samples in which victims are overrepresented.<sup>2</sup> A fundamental problem

---

<sup>1</sup>Throughout the paper we use the term domestic violence to indicate violence by the significant male other. The latter is sometimes also denoted by intimate partner violence.

<sup>2</sup>Using a small sample of victims, Tauchen, Witte and Long (1991) find that the severity of domestic violence against women decreases with the fraction of the year their male partner is employed. Employment of the female does not affect the severity of domestic violence on average; however, for our purposes it is useful to point out the finding that in low income working families, higher female

of studies based on survey data is that in surveys victims may underreport the actual exposure to violence. With domestic violence in particular, women may feel embarrassed disclosing victimization or they may want to protect the perpetrator. Naturally, surveys can only record retrospective information, and cognitive dissonance may even cause victimized respondents to forget exposure to violence. An additional problem is that different groups of individuals may apply different definitions of what constitutes violence. In case of a systematic difference between unemployed and employed individuals, this could bias the estimates of the association between labor market state and victimization towards zero. With police records or crime records, these issues are echoed. Indeed, in the case of domestic violence, the initiation of legal proceedings has dramatic consequences for the future of the relationship between perpetrator and victim, so that only the worst cases can be expected to be recorded.

In this paper we follow a different approach to observe violence, by using medical register data. Specifically, we use a unique set of merged population register data from the province of Skåne (i.e., South Sweden) covering the period 1999-2008, containing merged information at the individual level from the population register, the income tax register, and the in-patient and out-patient registers. The population register and income tax register capture individuals' labor market status. The in-patient register contains all contacts with the medical sector that lead to overnight hospital stays. The out-patient register contains all ambulatory care contacts including all contacts (visits and telephone calls) with physicians. Diagnoses are recorded for each contact. Contacts are recorded by the hour. The diagnoses are expressed at the highest level of detail of the Swedish dialect of the ICD classification system (version 10). In this system, injuries due to external causes are accompanied by codes that provide information on the specific cause of the injury, including various types of interpersonal violence. In addition, these codes provide categorical information on the location of the offence. This enables the construction of indicators of victimization of (domestic) assault at the individual level. The diagnoses also contain supplementary indicators for the oc-

---

employment is associated with lower severity of domestic violence. Bowlus and Seitz (2006), using a survey among a random sample of women, find that employed women are less likely to be abused by their husband. FRA (2014) finds that victimization due to violence is more likely among unemployed women than among employed women or non-participants in the labor market, in the EU. A number of other studies use data from developing countries. Choi and Ting (2008) use data from the 1998 South African Demographic and Health Survey to analyze how the distribution of income across spouses affects domestic violence. The results are mixed. Chin (2012), using data from India, finds no evidence that female labor force participation is associated with a lower incidence of violence against the woman. Grogan and Sadanand (2014) find a positive association between labor force participation and violence victimization in Latin America, but they argue that this positive sign may disappear once women become sufficiently wealthy.

currence of abuse and for having relationship problems that result in an atmosphere of violence and lack of control. In total, these indicators enable us to obtain a rather comprehensive account of the extent of victimization due to (domestic) violence.

Medical and health care registers have a number of advantages as a source of information on the occurrence of violence. First, they exploit the medical necessity to contact a health care worker after violence, maltreatment or abuse.<sup>3</sup> Not seeking help may have serious health repercussions. Secondly, the diagnoses are reported objectively by health care workers using the ICD classification system. These workers are experienced in this reporting as they typically perform this many times per working day. Thirdly, reporting violence does not have repercussions coming from the perpetrator or the social environment. This is because health care workers are tied to the Hippocratic Oath and the Declaration of Geneva which require any disclosure to remain confidential.<sup>4</sup> Fourthly, recall issues are not relevant. Fifthly, the registers cover the full population. And sixthly, the registers provide direct cost measures of treatments.

A few other studies have used medical registers to observe economic determinants of exposure to violence (see Estrada, 2006, Aizer, 2010, Kruse et al., 2010, and Curca, Dermengiu and Hostiuc, 2012). These studies only had access to in-patient registers or hospital registers. As we shall see below, the vast majority of cases of violence in our data are in fact recorded in the out-patient registers, notably those involving contacts with physicians. Furthermore, these studies do not focus on victimization among young females, with the exception of Kruse et al. (2010) who find that, among women below 30, “not working” is a risk factor for violence victimization. They use the hospital emergency room patient register to observe violence and they do not distinguish between domestic and non-domestic violence.

To avoid misunderstandings we should emphasize that the empirical analysis does not aim at the identification and estimation of causal effects of unemployment on victimization due to violence. Any association between exposure to violence and the victim’s labor market status may simply reflect selectivity. A causal analysis requires different econometric methods and assumptions than those adopted in this paper. Instead, we aim at establishing associations. These are useful for policy purposes because they indicate in which social strata violence is concentrated. For example, if an association is established then public policy may target a domestic-violence awareness

---

<sup>3</sup>FRA (2014) reports that female victims of violence are more likely to consult or be in contact with health services than with any other professional organization or agency. Kruse et al. (2010), which, to our knowledge, is the only study to date that compares individual violence victimization records across different types of population registers, finds that many more cases of violence are recorded in the hospital emergency room patient register than in the crime victim register in Denmark, 2002-2006.

<sup>4</sup>The one exception concerns children. Health care workers are obliged to contact the authorities if child abuse is suspected.

campaign towards neighborhoods with high youth unemployment.

The outline of the paper is as follows. In Section 2 we present the data including the medical registers. We construct our sample of female youths aged between 16 and 25 by combining data from several population-wide registers including administrative health care registers and the annual income tax register. We describe these data and the variables used in detail. In particular, Section 2 includes a careful description of how we construct different violence indicators. Section 3 presents and discusses the estimation results. Section 4 concludes.

## 2 The data

### 2.1 The health care registers and the Swedish health care system

We use population-wide registers of records at the individual level. The individual indicators of exposure to domestic violence are derived from individual health care records in the 1999–2008 “patient administrative register systems” PASiS and PRIVA from the region of Skåne in South Sweden. These two registers are administrated by the Regional Council of Skåne and contain detailed records of all occurrences of in-patient and out-patient care for all inhabitants of the region. Here, “in-patient” refers to visits or spells at medical units that include at least one night’s stay. These are mostly overnight hospital treatments. “Out-patient” refers to all other contacts with care providers, i.e., all ambulatory care, such as day-time visits to physicians, dentists, therapists, emergency care units, specialized nurses, and physiotherapists. In addition, it covers consultations by telephone.

To understand how these registers are created it is useful to outline the Swedish health care system in the years covering our observation interval. Health care is mostly public, organized at the county level. Within a county (such as Skåne), different communities have different health care centers that house all outpatient care. Typically, a rural community has only one center. Larger cities have multiple centers. Every individual is assigned to exactly one health care center. This is usually the nearest center. Each center has a team of physicians, first-aid workers, and nurses. In case of a need to see a health care worker, including first-aid and emergency aid, an individual goes to the center and is helped by the next available appropriate health care worker. There is no path dependence in the identity of the health care worker across consecutive contacts. For a given contact reason, on a given day, incoming individuals are dealt with sequentially by the first available health care workers. Workers in the health care sector (from nurses to hospital specialists) are county civil servants. The health care

system is funded through a proportional county tax on income. Health care usage is free, with the exception of a small deductible which in our observation window is capped at about 80 euro per adult person per year.

A small number of health care providers (notably dentists) are private. The PASiS register contains all publicly provided care, whereas PRIVA contains all privately provided care. The information in PASiS and PRIVA includes dates of admission and discharges, as well as detailed diagnoses and DRG-based costs.<sup>5</sup>

As explained in Section 1, the occurrence of an act of violence is observed from the specific diagnoses as described in the ICD-10 codes that are recorded at each contact with the health care provider. These codes are at the four digit (or rather one letter and three digits) level. Many unique health care contacts are captured by multiple codes capturing different aspects of the health problem. In Subsection 2.3 we discuss this in detail.

At the county level, health care registers are collected because they determine the monetary streams from the county to the various health care centers and hospitals. At the same time, at the national level, the register data are collected as part of the so-called “National eHealth” endeavor to improve efficiency in health care. Here, institutional variation in the health care systems across counties is used for “natural experiments” in the analysis of the connection between health care diagnoses and treatments and health outcomes. For this reason, the national health authorities place great value in the collection of reliable health-care diagnosis records.

In many countries in the world, individuals have a personal physician, and this is usually also the physician of the household members, including, in the case of domestic violence, a possible perpetrator. For our purposes, the absence in Sweden of such a personal physician is an advantage, as it further reduces the likelihood that the perpetrator is informed about the information flow originating with the victim, as well as the likelihood that he is held accountable for the act of violence.

## **2.2 Unemployment status and personal characteristics**

We now turn to the data on individual unemployment and personal characteristics. In Sweden, each individual has a unique identifier which is used to record all contacts with the health care system as well as the general public administration, tax boards, employment offices and so on. We use this to match the above-mentioned health care registers to individual information on socio-economic and demographic conditions. Specifically, we merge the health care registers to a dataset that itself consists of a

---

<sup>5</sup>The studies by Kristensson, Hallberg and Jakobsson (2007) and Nilsson and Paul (2014) also use these registers.

number of different registers. This dataset has been used before by Meghir and Palme (2005) and covers all persons born in Sweden between 1940 and 1985, their parents, and all their children. It includes variables from the annual LISA register on incomes by type, work absence days, detailed education measures, as well as information on date of birth, marital status, vertical family connections across different individuals, and migration status. This dataset is annual, in the sense that each variable is only recorded once a year. It covers the years 1992–2002 and 2004–2006.<sup>6</sup> The “Skåne dataset” that forms the intersection of the health care registers and the Meghir and Palme (2005) dataset contains about 1 million individuals, which is the vast majority of inhabitants of Skåne in 1999–2008. Note that, in terms of calendar time, the intersection of the health care registers on the one hand and the socio-economic / demographic registers on the other constitutes seven years: 1999–2002 and 2004–2006.

To capture individual unemployment in a year  $t$  we have a limited number of variables at our disposal. First, we observe whether the individual is employed in November of year  $t$ . Secondly, we observe total annual income from labor and the total amounts of sickness absence benefits, parental leave benefits, disability benefits, and unemployment benefits, received in  $t$ .

Accordingly, we define an individual to be unemployed in year  $t$  if one of the following two conditions applies. First, the individual receives no labor income, sickness absence benefits, disability benefits or parental leave benefits in  $t$  but does receive unemployment benefits in  $t$ . Secondly, the individual is not employed in November but receives labor income, sickness absence benefits, disability benefits or parental leave benefits during the year.

Notice that the definition does not include individuals who do not receive any income from labor, sickness absence benefits, parental leave benefits, disability benefits or unemployment benefits in year  $t$ . Young individuals without observed income are likely to be students. Unfortunately, our data do not provide direct observation of student status. Therefore, for each year  $t$ , we remove all individuals without any of the above income types in  $t$  from the sample, because it is not possible to state whether they are unemployed without any income or whether they are a student. That is, we remove these individuals for year  $t$  if this income condition applies in  $t$ . In fact, most students in Skåne reside in the communities of Malmö and Lund. Therefore, in a sensitivity analysis, we do not remove those individuals in  $t$  who are without any income in  $t$  and who do not live in Malmö or Lund in  $t$ .

The LISA register with the income and employment variables only contains individ-

---

<sup>6</sup>The LISA registers for the years 2007 and 2008 were not available at the time at which we applied for and received the data. Variables from the LISA register for the year 2003 are not provided to us. See SCB (2009) for a detailed description of the variables in the LISA register.

uals aged 16 and above. Further, our focus on female youths leads us to exclude women aged above 25. As a result, the sample of women in  $t$  only contains women aged in the interval 16–25. Recall that we observe violence outcomes and unemployment status at the individual level for at most 7 years per individual. The above subpopulation selection procedure results in a total of 328,038 woman-year combinations. Obviously, most women appear more than once in this number. Indeed, the sample contains 104,316 women, so we have on average 3.14 years of observations per woman. Of these women, 5479 are observed in all 7 years. Many women are only observed in the first few years or the last few years because they cross the age thresholds 16 and 25 at some point within the interior of the calendar time window.

As it turns out, among those who satisfy the above unemployment definition in year  $t$ , the vast majority is not employed in November but does receive labor income during the year. It is likely that some of these are in fact students who were employed in a temporary job during the summer. Unfortunately, our data do not allow us to isolate such a subset. As a result, the annual “unemployment indicator” that follows from the above unemployment definition merges job seekers with individuals who only worked part of the year and were non-participant in the labor market during the remainder of the year. It turns out to be positive for half of the woman-year combinations. As such, the “unemployment indicator” is not a conventional unemployment variable. Rather, it captures whether an individual has a loose connection to the labor market or a strong connection to the labor market. Here, “loose” and “strong” can be replaced somewhat informally by “job seeking or temporarily working” and “regularly employed most of the time”, respectively. The association between this indicator and violence victimization then captures the association between distance to regular work on the one hand and violence victimization on the other hand.

For obvious reasons it is useful to have information on the existence of a male partner who is in a steady relationship with the individual of interest, and on the unemployment status of such a partner. However, partnership status is not well-observed in the data.<sup>7</sup> In the next subsection we show that some of the violence indicators are

---

<sup>7</sup>Joint household earnings are only provided for individuals in couples living in the same household who are married or who have joint children. From this, one may identify a man-woman couple in a given year  $t$  if the following conditions are all satisfied: *(i)* they live in the same parish in year  $t$ , *(ii)* they have the same family type in  $t$ , and their children have the same ages, *(iii)* the gross annual joint household earnings in  $t$  are identical for the man and the woman and are nonzero, and *(iv)* the man and the woman are the oldest two household members. Not surprisingly, among the woman-year combinations in the sample, only a small fraction is identified as being part of a such couple. Presumably, many women below 25 are single or have a partner who lives elsewhere or they cohabit without children, for most years before age 26. In sum, it is likely that the data at our disposal do not enable us to identify many couples that do exist in reality.

nevertheless informative on whether violence is domestic or not.

## 2.3 Measures of violence

As noted in Section 1, we extract a number of violence indicators from the health care registers. Our first measure follows Aizer (2010) by focusing on assaults. We measure assaults in our data by way of diagnoses coded in the ICD-10 international classification of diseases. Each ICD-10 code consists of a capital letter denoting broad categories or “sections”, followed by a number of digits. The sections S and T refer to injuries, poisoning and other health consequences of external causes. If a code in these sections is recorded to diagnose an injury due to an external cause, then supplementary codes from the V, W, X and Y sections are used simultaneously to characterize the external cause in terms of the events or circumstances that led to the injury. In these sections, the codes X85-Y05 and Y08 and Y09 capture assaults as the cause of the injury. In this set of codes, the by far most common categories are Y04 (assault by bodily force) and Y05 (sexual assault). Y04 includes unarmed brawls and fights but not fights involving the use of a weapon or strangulation or rape. Y05 includes rape. The other assault codes include e.g. strangulation, suffocation, drowning, usage of a handgun, or a firearm, or an explosive, or a hot object, or a sharp object, such as a knife, or a blunt object, pushing, usage of a motor vehicle, and usage of a baseball bat.

Recall that, in contrast to Aizer (2010), we observe all health care visits related to assaults and not just female hospitalizations for assaults. Thus, our measure is more comprehensive and includes, for example, visits to a family doctor due to assault-related injuries. Indeed, over 90% of all observed assaults in our data are from the out-patient register and are thus not recorded in the hospitalization (in-patient) register. We should note, however, that hospitalization in Sweden by definition includes at least one overnight stay in the hospital, whereas this is not the case in certain other data such as Aizer’s (2010) data which are from California. Our in-patient register is therefore more selective than hers.

In the Swedish ICD-10 dialect, the assault codes contain an additional digit specifying the place of occurrence of the assault. Specifically, if the additional digit is zero then the assault took place at home. Other values capture assault on the street, at school, in a service area, in institutional housing, in a sport facility, at a farm, in an industry area, and somewhere else / unspecified. We use this information to create a measure of domestic violence. Clearly, although “domestic” literally means “at home”, violence at home is not synonymous to domestic violence in the sense of intimate partner violence. Unfortunately, the ICD-10 assault codes in our data are not accompanied by information on the nature of the perpetrator, so that we cannot replace the infor-

mation on the place of occurrence by information on the perpetrator. It has been argued that most violence against women is actually by the intimate partner (Aizer, 2010) so that assaults at home may under-estimate total intimate partner violence. Indeed, some of the violence reported to have occurred outside of the home may in fact be intimate partner violence. Among female youths, assaults at home only constitute 14% of all assaults (see below). All this suggests that assaults at home are a rather tight measure of intimate partner violence. Moreover, we conjecture that some of this violence is committed by boyfriends or (ex-)partners not living in the same household.<sup>8</sup>

As an alternative measure of intimate partner violence we use a code from the Z section of ICD-10. The codes in this section are not used to diagnose diseases, illnesses or injuries. Instead, they capture contacts with health care workers about non-medical causes of an illness or injury and/or about problems that are relevant for the individual's health. They include codes for unemployment and poverty but these are not used in practice. For our purposes, the code Z630 is particularly interesting. It captures problems in the relation of the individual with her spouse or partner; more specifically, discord and hostility in the relationship resulting in a loss of control and an atmosphere of violence. The violence could be hitting or striking the partner. A typical context in which this code is used is when a woman mentions relationship problems along these lines during a visit to a physician, as supplementary information for current or future health problems or in order to ask for advice on how to prevent further escalation or for advice on treatments such as therapies. Clearly, this measure focuses on domestic, i.e. partner-related, problems. Since an atmosphere of violence is explicitly mentioned, the Z630 code seems to be a good proxy of domestic violence.

The assault measures capture acute domestic and non-domestic violence. We complement these measures with a measure of prolonged (i.e. less acute) abuse. This measure corresponds to ICD-10 code T74 which captures maltreatment syndromes that have been confirmed in the sense that more evidence is available than just a suspicion on behalf of the health care worker. Examples are neglect, the battered spouse syndrome, and physical, sexual or psychological abuse. This type of violence is not necessarily acute but may have been prevalent for a long time and is regarded to be very serious. Importantly, by design, it is not used in conjunction with the "violent atmosphere" measure in given contacts with health care workers. Also, it is not exclusive for domestic violence as it is not ruled out that the perpetrator is a different family member, neighbor, colleague or other person.

---

<sup>8</sup>Alternatively, one may conjecture that some of the assaults against young women at home involve the father as perpetrator. In Tertilt and Van den Berg (2013) we examine this in detail, using information on the family type. We conclude that assaults by the father are only of minor importance in the analysis of assaults at home.

To sum up, we have four measures of violence. Assaults capture acute over-all violence. Assaults at home are a subset of assaults and capture acute domestic violence. “Violent atmosphere” captures domestic violence. And “Maltreatment syndrome” captures prolonged abuse whether domestic or not. We believe that together these measures provide a comprehensive picture of victimization due to violence against females. Table 1 lists the total number of observed cases. These concern all women who are aged between 14 and 25 and regardless of whether their labor market status can be observed.

Table 1: Occurrences of violence against young women in Skåne, 1999–2008

Variable	ICD codes	# cases
Assaults	X85–X99, Y00–Y05, Y08–Y09	1867
Assaults at home	same as above but with extra digit 0	266
Violent atmosphere	Z630	777
Maltreatment syndrome	T74	2633

To put the importance of violence against female youths in perspective, it is important to emphasize that a large fraction of assaults against women in the population concerns young women. The numbers displayed in Table 1 for assaults, assaults at home, violent atmosphere and maltreatment syndrome amount to 49%, 41%, 30% and 52% of the corresponding numbers for all women aged 14 and above.

In the empirical analysis in this paper we use outcome variables that are derived from the above measures. First of all, we take binary exposure variables as the individual yearly outcomes. This serves to prevent that multiple visits per woman per year have large effects on the results. This is particularly relevant for the “violent atmosphere” measure where the number of victims is about 30% of the number of contacts. As a result, the number of woman-year combinations for whom an outcome is positive is lower than in Table 1. In addition, we require that the woman’s unemployment status is observed (see Subsection 2.2), and for this, the woman needs to be at least age 16, the years 2003, 2007 and 2008 are omitted, and the woman should receive at least one of the income types that we observe. As noted in the previous subsection, this results in a set of 104,316 women and 328,038 woman-year combinations to be used in the empirical analysis in Section 3 below. In this set, the binary annual outcome variables corresponding to assaults, assaults at home, violent atmosphere and maltreatment syndrome are positive for 397, 51, 104 and 469 cases, respectively.

## 3 Results

### 3.1 Parameter estimates

To specify the models we estimate, let the index  $i$  denote an individual, and let  $Z_{i,t}$  denote the unemployment status of the woman in year  $t$ . Further, let  $X_{i,t}$  denote personal characteristics (age, highest current level of education, migrant status, community, calendar time) and let  $Y_{i,t}$  be the relevant outcome variable for the individual occurrence of victimization. The regression of  $Y$  on  $Z, X$  can then be specified by way of a linear probability model,

$$Y_{i,t} = \beta_0 + \beta_1 Z_{i,t} + \beta_2 X_{i,t} + \varepsilon_{i,t} \quad (1)$$

We estimate separate models for each of the four outcome variables. Concerning the covariates we point out that of all woman-year combinations, 3% concern immigrants, and 13%, 48% and 39% have low, intermediate and high education, respectively. Note that some women are still in education in some years  $t$  so that the highest currently attained level of education is a lower bound for the obtained amount of education. Per outcome variable we lose a small number of individuals due to missing covariates. This results in a loss of up to 6 women with a positive binary annual outcome, which we believe is negligible.

Table 2 reports the estimates. We cluster the standard errors at the individual level (i.e., at  $i$ ), since many women in the sample provide multiple annual observations.

In linear probability models, coefficients are usually interpreted as covariate effects in percentage points on the probability that the outcome is true (i.e., has the value 1). In our setting, this means that the unemployment coefficient is the percentage-point change in the annual probability of victimization associated with a change in the binary unemployment status. For example, in the case of all assaults, the annual probability of victimization increases by 0.31/1000 if the value of the unemployment status variable changes from 0 to 1. In relative terms, this means that the annual probability of victimization increases by 26% (this follows by comparing 0.31/1000 to 397/328038). This is a substantial increase. In aggregate absolute terms it translates into 102 young women per year. In the data, the unemployment indicator equals 1 for about half of the woman-year combinations. If those women switch from unemployment to employment then the number of victims of at least one assault in Skåne in a given year would be about 51 lower. To this one may add the change in the number of victims with a maltreatment syndrome.

As a first sensitivity analysis, we estimate models that exclude calendar time and community indicators and that do not allow for clustered standard errors. The coef-

Table 2: Annual probability of violence victimization, among female youths.

characteristic	All assaults	Assaults at home	Violent atmosphere	Maltreatment syndrome
log(age-12)	-0.01	0.01	0.19	0.14
unemployed	0.31**	-0.02	0.09	0.47***
immigrated	0.01	0.06	0.58*	-0.09
education: intermediate	-2.48***	-0.34***	-0.41**	-3.12***
high	-3.32***	-0.48***	-0.60***	-4.03***
calendar year	0.18***	0.02***	0.01	0.21***
constant	-350 ***	-47.3***	-24.7	-425***

Note: estimates are coefficients\*1000 in linear probability models. The models control for community (33 binary indicators). Standard errors are clustered at the level of the woman. The superscripts \*\*\*, \*\*, and \* refer to two-sided significance at the 1%, 5% and 10% level, respectively. # women = 104,316; # woman-year combinations = 328,038.

ficients are qualitatively and quantitatively very close to those in Table 2. Next, we estimate the associations by way of probit models,

$$\Pr(Y_{i,t} > 0) = \Phi[\alpha_0 + \alpha_1 Z_{i,t} + \alpha_2 X_{i,t}] \quad (2)$$

where  $\Phi$  is the cumulative standard normal distribution function. The results (not shown here) are qualitatively identical to those based on the linear probability models.

Two additional sensitivity analyses address the extent to which our individual unemployment indicator captures summer jobs among women in full-time education (in addition to capturing the strength of the connection to the labor market among more conventional labor force participants). First, we include women in  $t$  who are without any income in  $t$  provided they do not live in Malmö or Lund in  $t$ . We define such women to be unemployed in  $t$ . Recall that women aged below 26 not living in Malmö or Lund are less likely to be students than those living in Malmö or Lund. This increases the sample size with 7557 women and 54,038 woman-year combinations, to a grand total of 111,873 women and 382,076 woman-year combinations. It turns out that the estimation results (not shown here) are very similar to those reported above. In particular, the coefficients on the unemployment indicator and their standard errors are virtually identical to those reported above.

Secondly, we further restrict the age window to 19–25. The rationale behind this is that women aged 16 to 18 are likely to be enrolled in full-time education. Among such women, a value of one of their unemployment indicator is more likely to be the

result of a summer job rather than a reflection of a large distance to attractive labor market opportunities. Table 3 reports the results.

Table 3: Annual probability of violence victimization, among females aged 19–25.

characteristic	All assaults	Assaults at home	Violent atmosphere	Maltreatment syndrome
log(age-12)	-0.92***	-0.04	0.12	-1.12***
unemployed	0.34**	0.00	0.13	0.61***
immigrated	-0.08	0.10	0.50	0.13
education: intermediate	-2.78***	-0.45**	-0.62**	-3.41***
high	-3.67***	-0.58***	-0.79***	-4.44***
calendar year	0.20***	0.02***	0.02	0.23***
constant	-389***	-48.4***	-47.3	-449***

Note: estimates are coefficients\*1000 in linear probability models. The models control for community (33 binary indicators). Standard errors are clustered at the level of the woman. The superscripts \*\*\*, \*\*, and \* refer to two-sided significance at the 1%, 5% and 10% level, respectively. # women = 89,428; # woman-year combinations = 252,471.

Clearly, the estimation results are very similar to those reported in Table 2, especially those on the unemployment indicator and their standard errors. Notice that some outcomes now strongly decrease with log(age-12). Perhaps the outcome does not vary with age until age 19, or perhaps the outcome among students and pupils is less age-dependent than among young labor force participants. An alternative explanation is that at higher ages the logarithmic transformation of age makes the outcome less responsive for a given coefficient.

In sum, the final two sensitivity analyses confirm that the unemployment indicator captures the strength of the connection to the labor market and that the coefficients of interest capture the associations between this strength and violence victimizations.

### 3.2 Discussion

The first main finding is that female unemployment is associated with a higher prevalence of violence victimization. This follows because for two of the four violence measures the association is significantly positive. Specifically, we find a positive association for the over-all assault measure and for the maltreatment syndrome measure.

We do not find an association for the “assaults at home” measure, as it does not vary with the unemployment status. To the extent that the difference between “assaults”

and “assaults at home” captures non-domestic violence, this finding suggests that the association between unemployment and assaults is mostly due to non-domestic violence. This in turn is consistent with the fact that there is no association between unemployment and a violent atmosphere at home. After all, the latter measure also captures domestic violence.

The positive association for the maltreatment syndrome measure may then also be attributed to non-domestic maltreatment. However, when interpreting the findings it is important to keep in mind that associations reflect selectivity of the unemployment status. That is, individuals with characteristics that lead to unemployment may also be more likely to be violence victims. Such a selectivity may be less important for assaults at home than for maltreatments at home. For example, assaults at home may be primarily driven by turbulence in society. Conversely, the occurrence of a maltreatment syndrome may reflect personality traits or lifestyle characteristics of the victim that affect the likelihood both of victimization and of unemployment. One may speculate about examples, such as adverse non-cognitive skills or having grown up in a poor neighborhood (see e.g. Rusbult and Van Lange, 2003, for an overview of the social-psychological literature). Such characteristics may on average lead to a low socio-economic class in adulthood, and through that to a partnership with an abusive man as well as to unemployment. The victimization and the unemployment may be particularly likely after a lengthy deterioration of the living conditions of the victim. In this context it is important to remember that the maltreatment may have started years before the observed diagnosis. The assaults measures and (to a lesser extent) the violent atmosphere measure capture more acute conditions.

Keeping this in mind, the results suggest that maltreatment and non-domestic assaults are more common among unemployed young women than among employed young women. Moreover, assaults at home and the occurrence of a violent atmosphere at home are not associated with the victim’s unemployment status. The unemployment status may then be used as a marker for the occurrence of maltreatment and non-domestic assaults. Policies aimed at preventing or monitoring non-domestic assaults and maltreatment of young women may then focus on unemployed young women, in order to increase the policy exposure. Alternatively, such policies may focus on neighborhoods with high female youth unemployment. FRA (2014) reports that in the EU, women who have experienced violence are more likely to say that they also know of other female victims. This suggests that victims are socially clustered in society. The female youth unemployment rate may be used to identify such clusters.

Having said this, the estimation results show that a woman’s level of education is a much stronger marker of violence victimization than her unemployment status. In the sample, 13% of the young women have the lowest level of education. The coefficient

capturing the difference in victimization compared to those having higher education levels is about 5 to 10 times as large as the coefficient capturing the difference in victimization that is associated with a change in the unemployment status. Moreover, the coefficients for education are strongly significant for all outcome measures, including assaults at home and the occurrence of a violent atmosphere at home. In sum, a low level of education is a much more informative marker for violence victimization than the unemployment status. Consequently, if the individual level of education and the unemployment status are equally costly to observe in practice, then policies aimed at preventing or monitoring violence against young women should first identify the low-educated women and only after that focus on the unemployed. Of course, policy makers with access to the registers we use observe both variables at no cost. However, recall that the unemployment status variable in the register is an imperfect indicator of (involuntary) unemployment as defined by e.g. the ILO. It is an open question whether conventional measures of unemployment have more explanatory power for exposure to violence among young women than the indicator we use.

We may relate the estimated effect of education to the discussion on causal and selection effects. Education has a negative effect on unemployment, so education creates a positive association between unemployment and violence victimization. This means that if we omit education from the analysis then the resulting association between unemployment and violence victimization is partly due to the effects of education on these variables. This part of the total association then reflects a selection effect. In other words, by controlling for education we ensure that the estimated association is not due to this selection effect.

## 4 Concluding remarks

Using novel health care register data from Sweden, we construct four indicators of violence exposure. We merge these health care data with unemployment and other personal characteristics from other registers, and we examine the association between violence exposure and unemployment, among young women. We find that female victimization is more prevalent amongst unemployed women compared to their employed counterparts. This finding holds only for measures of general violence (assaults and maltreatment) that do not necessarily capture domestic violence. We do not find a significant association between specific measures of domestic violence (assaults at home, and problems in a relationship resulting in a violent atmosphere) and unemployment.

The unemployment status may be used as a marker for the occurrence of maltreatment and non-domestic assaults. Policies aimed at preventing or monitoring non-domestic assaults and maltreatment of young women may then focus on unemployed

young women, in order to increase the policy exposure (i.e., to reach potential victims and perpetrators).

It is important to repeat the point that the associations that we detect cannot be translated into causal effects of the woman's unemployment status on violence victimization. The associations we measure are produced both by causal effects and by selectivity on unobservables. Moreover, the associations may reflect reverse causality from victimization to the unemployment status as well, since it cannot be ruled out that the victimization occurs early in the calendar year and the job loss occurs later in that year.

Obviously, it is important to know whether a young woman's unemployment has a causal effect on the probability of violence victimization. A positive causal effect implies that fighting unemployment among young women has the beneficial effect of causing a reduction of the probability that they are victims of violence. The results in the present paper are not informative on this, so they cannot be used as a justification to fight unemployment among young women. More in general, it is an important challenge to increase our knowledge about causal economic determinants of violence and the causal pathways originating from these determinants. Such knowledge should have a more universal relevance than results from descriptive studies. Again, this requires causal inference. Our companion paper Tertilt and Van den Berg (2013) addresses these issues in detail.

If our results are driven by selectivity, with violence being clustered in certain neighborhoods, and with women with poor labor market qualifications (and hence higher unemployment) being more likely to live in such neighborhoods, then this strengthens the case for prevention policies focused on specific neighborhoods. If instead there is a causal component, then, rather than targeting regions, it would be better to focus prevention policies on young women who just lost their jobs. Or, possibly, no separate prevention policies might be necessary as general unemployment reduction policies might have the positive side effect of reducing violence against young women as well.

One would also like to better understand why the association holds for general violence but not domestic violence. Clearly the mechanisms need not be the same. Domestic violence is related to frictions with the intimate partner and perhaps also with the woman's outside option (or lack thereof) – issues that do not apply to other types of violence. Conversely, general violence may be related to an atmosphere of fear and aggression in a neighborhood, a setting to which families perhaps provide some insurance. Clearly these are pure speculations and further research beyond the current paper is necessary before reaching such conclusions.

## References

- Aizer, A. (2010), “The gender wage gap and domestic violence”, *American Economic Review*, 100(4), pp. 1847–1859.
- Bowlus, A.J. and S. Seitz (2006), “Domestic violence, employment, and divorce”, *International Economic Review*, 47(4), pp. 1113–1149.
- Chin, Y.M. (2012), “Male backlash, bargaining, or exposure reduction?: women’s working status and physical spousal violence in India”, *Journal of Population Economics*, 25, 175–200.
- Choi, S.Y.P. and K.F. Ting (2008), “Wife beating in South Africa”, *Journal of Interpersonal Violence*, 23(6), pp.834–852.
- Curca, G.C., D. Dermengiu and S. Hostiuc (2012), “Patterns of injuries in domestic violence in a Romanian population”, *Journal of Interpersonal Violence*, 27, 2889–2902.
- Estrada, F. (2006), “Trends in violence in Scandinavia according to different indicators - An exemplification of the value of Swedish hospital data”, *British Journal of Criminology*, 46, 486–504.
- FRA (2014), “Violence against women: an EU-wide survey: main results”, European Union Agency for Fundamental Rights, Vienna.
- Fougère, D., F. Kramarz, and J. Pouget (2009), “Youth unemployment and crime in France”, *Journal of the European Economic Association*, 7(5), 909–938.
- Grogan, L. and A. Sadanand (2014), “Female employment and violence in the household: Evidence from Nicaragua”, Working paper, University of Guelph.
- Kristensson, J., I.R. Hallberg, and U. Jakobsson (2007), “Healthcare consumption in men and women aged 65 and above in the two years preceding decision about long-term municipal care”, *Health & Social Care in the Community*, 15, 474–485.
- Kruse, M., J. Sørensen, H. Brønnum-Hansen and K. Helweg-Larsen (2010), “Identifying victims of violence using register-based data”, *Scandinavian Journal of Public Health*, 38, 611–617.
- Lloyd, S. (1997), “The effects of domestic violence on women’s employment”, *Law and Policy*, 19(2), 139–167.
- Meghir, C. and M. Palme (2005), “Educational reform, ability and parental background”, *American Economic Review*, 95, 414–424.
- Nilsson, A. and A. Paul (2014), “The effect of copayments on children’s and adolescents’ demand for medical care”, Working paper, Lund University.

- Rusbult, C.E. and P.A.M. Van Lange (2003), “Interdependence, interaction, and relationships”, *Annual Review of Psychology*, 54, 351–375.
- SCB (2009), Integrated database for labour market research, *Background Facts Labour and Education Statistics* 2009:1, Statistics Sweden, Örebro.
- Tauchen, H.V., A.D. Witte, and S.K. Long (1991), “Domestic violence: A non-random affair”, *International Economic Review*, 32(2), 491–511.
- Tertilt, M. and Van den Berg, G.J. (2013), “Domestic Violence over the Business Cycle”, University of Mannheim, unpublished manuscript.