OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 79, 3 (2017) 0305–9049
doi: 10.1111/obes.12177

A Cautionary Tale About Control Variables in IV Estimation*

Eva Deuchert† and Martin Huber†

†*Department of Economics, University of Fribourg, Bd. de Pérolles 90 1700, Fribourg, Switzerland (e-mail: martin.huber@unifr.ch)*

Abstract

Many instrumental variable (IV) regressions include control variables to justify (conditional) independence of the instrument and the potential outcomes. The plausibility of conditional IV independence crucially depends on the timing when the control variables are determined. This paper works through different IV models and discusses the (conditions for the) satisfaction of conditional IV independence depending on whether the control variables are measured prior to or after instrument assignment. To illustrate the identification issues, we consider the Vietnam War draft risk as instrument either for veteran status or education to evaluate the effects of these variables on labour market and health outcomes. Our empirical analysis based on the 'Young Men in High School and Beyond' survey suggests that commonly used conditional IV strategies to estimate the impact of draft induced military service or education may be severely biased due to the use of improper controls.

I. Introduction

The evaluation of causal effects in empirical economics frequently relies on instrumental variable (IV) methods. Identification requires that the instrument is associated with the endogenous treatment variable (IV relevance), but is not directly related with the outcome or unobservables affecting the outcome (IV validity). In many empirical applications, the latter assumption occurs only reasonable when controlling for further covariates, often referred to as conditional IV independence. This paper systemically works through different IV models and discusses the conditions for the satisfaction of conditional IV independence depending on whether the control variables are measured prior to or after instrument assignment.

It is remarkable that in many IV studies, the discussion and justification of conditional IV independence does not pay much attention to the time period in which the control variables are measured, i.e. whether this happens prior to instrument assignment or at a

JEL Classification numbers: C26, J24

*We have benefitted from comments by Dionissi Aliprantis, Per Johansson and seminar participants in Rotterdam, Aarhus, Linz and Bludenz.

later point. In particular, there seems to exist a wide spread consensus that it is reasonable to use IV methods in cross-sectional data, where outcomes and controls are measured in the same period. This stands in stark contrast with the programme evaluation literature relying on conditional independence of the treatment given observed controls, where it is well acknowledged that credible controls need to be measured prior to treatment assignment. Otherwise, they might be affected by the treatment so that conditioning on them likely introduces selection bias (e.g. Rosenbaum, 1984; Lechner, 2008). An intuitive discussion on this issue is provided in Angrist and Pischke (2009), who refer to such post-treatment variables that cause selection bias when being controlled for as bad controls.

There are typically two motivations for adding covariates: the instrument is not entirely random and associated with important confounders, or the instrument affects more than one treatment variable that is associated with the outcome of interest. In this paper, we use graphical tools (so-called causal diagrams, see Pearl, 1995) to discuss various IV models that postulate different causal relations between pre-instrument covariates, the instrument, the treatment, post-instrument covariates, and the outcome. For each scenario, we demonstrate if causal effects may be identified by IV methods when conditioning on pre-instrument covariates, post-instrument covariates, or both. This framework allows us to show that the timing of the determination and measurement of time-varying controls – albeit often ignored in the literature – crucially affects the plausibility of conditional IV independence and provides a guideline for selecting control variables in conditional IV frameworks.

We use the well-known Vietnam War lottery as an illustrative example for different conditional instrumental variable strategies. This is an ideal example for three reasons: First, the draft lottery is well-known instruments used in the literature (e.g. Angrist, 1990; Angrist, Chen and Frandsen, 2010; Angrist and Chen, 2011). Second, because the lottery was manually executed and capsules were inadequatly mixed, at least the first lottery failed and was not fully random. Draft numbers were associated with the month of birth (Fienberg, 1971) – an important covariate for later labour market, education or health outcomes (Buckles and Hungerman, 2013). And finally, the draft may have affected more than one treatment variable since it may have induced college enrolment due to college deferments (Card and Lemieux, 2001). We reconsider this instrument to discuss under which conditions causal effects can be estimated using conditional IV strategies.

In our empirical application we use data from the ‘Young Men in High School and Beyond’ survey that were collected shortly before and after individuals were at risk to be drafted. Our empirical analysis suggests that commonly used conditional IV strategies to estimate the impact of draft induced military service or education may be severely biased due to the use of improper controls. Firstly, using the draft lottery as an instrument for veteran status likely violates the exclusion restriction due to its effect on educational deferments. Secondly, controlling for veteran status when using the lottery as an instrument for education likely renders the instrument endogenous due to sample selection bias.

The issues discussed in this paper are not only relevant for the draft lottery, but occur in many empirical contexts. For instance, the problem that a supposed instrument affects several treatments simultaneously has also been noted by Lundberg (2005), who points out that variables alike 'gender composition of the first child/the first two children' have been used in different studies as instruments for 'fertility', 'marital stability', and 'parental time and resource allocation'. As a further example, Bazzi and Clemens (2013) mention that 'origin of a country's legal system' was used as an instrument for 'quality of economic institutions', 'inflation', 'inequality of land ownership', 'private sector credit', 'bank credit', 'stock market capitalization', 'average years of schooling in the population', and others.

The remainder of this paper is organized as follows. Section II describes the Vietnam draft and outlines how previous research used the Vietnam draft risk in conditional IV settings. Section III uses causal diagrams to define two common IV strategies: exploiting the Vietnam draft lottery as instrument for veteran status or for education. For each case, we discuss whether conditional IV independence holds and treatment effects can be identified when controlling for pre- and post-instrument covariates. We also show under which conditions none of these strategies identifies causal effects. Section IV provides empirical evidence suggesting that the assumption necessary to identify causal effects in these two scenarios are violated. The final section concludes.

II. Vietnam draft lottery instrument

During the Vietnam War, the majority of American troops consisted of volunteers, while the rest were selected through a draft (Gimbel and Booth, 1996). Young men at age 18 had to register at local draft boards for classification. These boards determined medical fitness and decided on the order in which registrants would be called. Men who had physical problems, were attending college, or were needed to support their families were granted deferments. Many men receiving deferments were from wealthy and educated families; most soldiers drafted were men from poor and working-class families.

In an attempt to make the draft more equal, a draft lottery was conducted in the years 1969–72 to determine the order of call to military service for men born between 1944 and 1952. The lottery assigned a draft number to each birth date for males in certain age cohorts, where low draft numbers were called first upon a ceiling. The first draft lottery in 1969 failed to be truly random – randomization was based on capsules that were inadequately mixed. Men with birthdates at the end of the year were more likely assigned low draft numbers (Fienberg, 1971).

The draft risk is frequently used as instrumental variable.[1] This literature can be classified into different strands. The first strand uses the draft risk as an instrument for veteran status (e.g. Angrist, 1990; Dobkin and Shabani, 2009; Angrist *et al.*, 2010; Angrist and Chen, 2011; Conley and Heerwig, 2011). A low random draft number (RDN) affects the likelihood to be an army veteran via two channels: through the risk of being drafted and through an increased likelihood to voluntarily join the army (Angrist, 1991) since the

[1] Angrist (1990), Conley and Heerwig (2011), Dobkin and Shabani (2009) or Lindo and Stoecker (2014) use an indicator for the lottery number being below the draft ceiling in a given year as instrument for Veteran status. Others group lottery numbers into several categories (Angrist and Krueger, 1992; Angrist *et al.*, 2010; Angrist and Chen, 2011). Note that making use of the draft lottery number requires information on the exact birth date of each respondent, which is often not provided in publicly available data sources for data protection reasons. A related instrument that does not rely on the exact birth date is the draft risk faced by each birth cohort (e.g. Grimard and Parent, 2007; de Walque, 2007), or by cohort and state (e.g. Malamud and Wozniak, 2010).

voluntary enlistment came with better terms, for example, through the possibility to choose the military branch and the timing of entry into service, as well as to accomplish special training or officer candidate programmes. Exploiting draft risk as instrument for veteran status assumes that the instrument affects the chosen outcome only via its impact on veteran status but has no direct effect, or any effect that runs through other variables. The second strand of literature uses draft risk as an instrument for education (Grimard and Parent, 2007; de Walque, 2007). This is motivated by the fact that education deferments were continued to be issued until 1971, which means that men could avoid being drafted by going to or staying in college (Card and Lemieux, 2001).

Finally, some authors instrument both education and veteran status at the same time. Angrist and Krueger (1992), for example, transform draft risk into several categories and instrument both endogenous variables with these categories. Malamud and Wozniak (2010) compute national and state level draft risks and use them as separate instruments. In both applications, however, the constructed instruments essentially come from one single phenomenon, namely draft risk, so that identification is delivered by functional form assumptions (which would not work in a completely flexible, nonparametric framework). An exception is Chaudhuri and Rose (2009), who use draft numbers as instrument for veteran status along with some additional instruments for education (i.e. geographical distance of the place of residence to a public or private four years college). For illustration, the appendix provides a graphical causal model that is compatible with their framework of multiple instruments. The availability of at least two instruments for the two endogenous variables solves the identification problem, but many applications may lack more than one credible instrument (if any).

III. Identification in different causal models

In this section, we consider the two IV strategies and demonstrate (non-)identification based on (model-free and thus, non-parametric) causal diagrams, while also providing a formal discussion based on equations. Causal diagrams (see e.g. Pearl, 1995) are often used in statistics, but have also gained popularity in economics (e.g. Chalak and White, 2011; White and Lu, 2011; Huber, 2014; de Luna and Johansson, 2014). They visualize a set of random variables and their causal effects by arrows and thus represent complex causal systems in an arguably more tractable way than long systems of equations for each of the random variables. This provides the researcher with a relatively intuitive way to understand which control variables are required for identification.

The key idea is to block confounding paths between variables, so that only the path of interest remains open. Two simple rules need to be considered when using causal graphs for identification: (i) In a causal path $A \to B \to C$, where each arrow represents the causal effect of one event (denoted by the capital letters A to C) on another, controlling for A, B or both blocks the path between A and C. It is not necessary to control for all variables to block a path, but it does in terms of identification not harm either. (ii) The direction of arrowheads is important. In a graph $A \to B \leftarrow C$, the variable B is referred to be a *collider*, defined as a common outcome of two variables It is necessary to carefully identify possible colliders in a graph because controlling for colliders creates a spurious association between to variables. In our example, controlling for B opens up a relationship between A and C

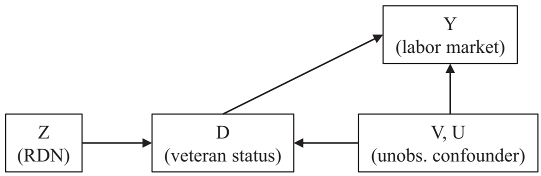


Figure 1 Causal diagram for the standard LATE framework

that was initially not there when not conditioning on B.[2] Note that collider bias, which is also known as sample selection bias (Heckman, 1979) or endogenous selection bias (Elwert and Winship, 2014), can occur no matter if one controls for a common outcome in a regression framework or stratifies on it (e.g. by selecting a sample based on a common outcome).

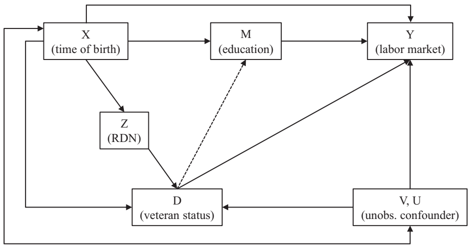
The standard IV model as a causal diagram

Figure 1 displays a standard IV model for our running example by means of a causal diagram, which consists of a treatment variable D (veteran status), an outcome Y (e.g. labour market performance or health), and unobserved terms U, V that affect the outcome and treatment, respectively, and may be arbitrarily associated with each other. The RDN Z serves as an instrument for veteran status. Identification is achieved since there is no direct causal connection between the instrument and the unobservables. Therefore, the effect of D on Y corresponds to the effect of Z on Y (intention to treat effect) divided by the effect of Z on D (first stage).

This simple framework, however, assumes that the instrument itself is not confounded and has no impact on the outcome other than through shifting the probability to have joined the army. In the following, we challenge these assumptions and show under which conditions a causal effect can still be identified and in which cases identification fails. We focus on two examples motivated by the previous literature: First, we show under which conditions the effect of veteran status can be identified when the instrument is not fully random and associated with the some pretreatment confounders (as in Angrist, 1990; Dobkin and Shabani, 2009). We also acknowledge that the treatment may affect more than one outcome (e.g. education, such as in Angrist and Chen, 2011).

Second, we discuss the usage of draft risk as an instrument for education. This is motivated by de Walque (2007) or Grimard and Parent (2007) who estimate the effect of college education on smoking. The authors of these studies acknowledge the possibility that the draft risk affects both college (which likely decreases smoking) and the probability of being a veteran (possibly increasing smoking through exposure to a smoking friendly environment and heavily subsidized tobacco products, Bedard and Deschênes, 2006). They therefore include veteran status as control variable to block the effect of the instrument on smoking operating through veteran status. We demonstrate under which conditions this

[2] As an example for collider bias, suppose the likelihood of entering college is positively affected by ability and negatively affected by the geographical distance to the next college. Assume that ability and place of living are statistically independent. In the subsample of individuals going to college, a person previously living far away from college is likely comparably able, as she joined college despite the geographical distance. In this case, conditioning on the collider variable college education creates a spurious positive association between ability and distance to college.



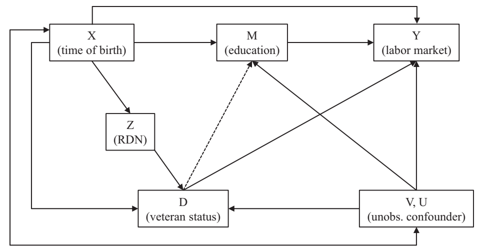


Figure 2 Causal diagrams satisfying conditional IV independence conditional on X

strategy identifies a causal effect. Finally, we show under which circumstances identification fails and why we believe that none of the mentioned IV strategies yields a causal effect.

Identifying the effect of veteran status when the instrument is confounded

Causal graph

As before, our causal diagrams consist of a treatment variable D (veteran status), an outcome Y (some labour market or health outcome), and the unobserved terms U, V that affect the treatment and outcome and may be arbitrarily associated with each other. The instrument Z (RDN) is confounded by a pre-instrument covariate X. In our example, this covariate is the timing of birth caused by the randomization failure of the first lottery. Many empirical applications therefore control for the month of birth to account for the fact that people born at the end of the year are different on key demographic characteristics than those born earlier in the year (season of birth effects, Buckles and Hungerman, 2013). Hence, pre-instrument covariates may not only affect draft risk, but also be directly associated with the outcome of interest or with unobserved variables. Moreover, we allow that veteran status has an impact on post-instrument variables (M). In our example, the GI bill finances continuous education after serving the military (Angrist and Chen, 2011). Consequently, education can be regarded as a mediating variable.

It is important to note that further unobservables that are not displayed in the causal graphs may enter the causal models, as long as they affect at most one of the elements (e.g. either X *or* Z), but not several jointly (e.g. not X and Z). As in Pearl (2000, p. 68), this allows treating these further unobservables as independent background factors, as they (by assumption) do not confound the causal links of any of our displayed elements.

In the left graph of Figure 2, we allow for a direct link of X on Y and/or an association between X and U, V, where the double-edged arrow between X and U, V implies that causation may go in either direction. When education is not affected by veteran status, the (total) causal effect of veteran status on the outcome can be identified when conditioning on X, or on both X, M. When education is affected by the veteran status, conditioning on pretreatment covariates identifies the total effect, while conditioning on (X, M) gives a partial direct effect.

In the right graph of Figure 2 we allow M to be confounded by the unobservables. As before, the total causal effect of veteran status on labour market or health outcomes is identified when conditioning on X. Controlling for both pre- and post-treatment covariates

(X, M) identifies only a causal effect if there is no association between veteran status and education. If the GI Bill induces higher education attainment after duty (so that there is an association between D and M), controlling for pre- and post-treatment covariates (X, M) does not identify any causal parameter – not even the direct effect. This is because in the right graph of Figure 2, education is a common outcome of veteran status and the unobserved variables. Controlling for education introduces an association between the instrument and U, V via D.[3]

Formal discussion based on a system of equations

For a formal discussion of the framework outlined in Figure 2, consider the following structural model that is for the sake of simplicity characterized by a system of linear equations with constant effects (albeit analogous arguments could also be made in nonparametric models at the cost of a more tedious notation).

$$Y = \alpha_D D + \alpha_M M + \alpha_X X + U, \quad (1)$$

$$M = \beta_D D + \beta_X X + W, \quad (2)$$

$$D = \gamma_Z Z + \gamma_X X + V, \quad (3)$$

$$Z = \delta_X X + Q. \quad (4)$$

The coefficients β, γ, δ represent the effects of the various variables in the system they are subscripted by and correspond to the causal arrows in Figure 1. U, V, W, Q are unobserved. While U and V may be arbitrarily correlated (thus causing the treatment to be endogenous), we throughout this paper impose independence of Q, implying that it is an independent background factor of Z in the sense of Pearl (2000):

Assumption 1. $Q \coprod (U, V, W, X)$, where $\coprod$ denotes statistical independence.

Concerning W, the assumptions that need to be imposed depend on whether the causal framework in the left or in the right graph of Figure 1 should be matched. Consider the following assumption:

Assumption 2. $W \coprod (U, V)|X$.

If Assumption 2 holds, the left scenario is satisfied, if W is associated with (U, V) even conditional on X e.g. when being a function of U and/or V, the scenario on the right holds. In either case, the probability limit of the simple IV estimator of the effect of D on Y given X, $\mathrm{Cov}(Y, Z|X)/\mathrm{Cov}(D, Z|X)$ coincides with the treatment's total causal effect. Path analysis reveals that the latter corresponds to $\alpha_D + \alpha_M \beta_D$, i.e. the sum of the direct causal path from D to Y, and the indirect effect of D on Y via M. IV identification is obtained because

[3] The same line of argument would apply when stratifying on M. This jeopardizes the usefulness of empirical results relying on draft risk as instrument in highly selected samples such as prison inmates (Lindo and Stoecker, 2014).

$$\begin{aligned}\text{Cov}(Y,Z|X) &= \text{Cov}(\alpha_D D + \alpha_M M + U, Z|X) = \text{Cov}(\alpha_D D + \alpha_M \beta_D D + \alpha_M W + U, Z|X) \\ &= \text{Cov}(\alpha_D D + \alpha_M \beta_D D, Z|X) = \text{Cov}(D, Z|X) \times (\alpha_D + \alpha_M \beta_D),\end{aligned} \tag{5}$$

where first and second equalities follow from substituting Y and M by their respective models after fixing X, the third from Assumption 1 and the last from rearranging terms.

In the left scenario of Figure 1, also the direct effect α_D is identified conditional on X and M by $\text{Cov}(Y,Z|X,M)/\text{Cov}(D,Z|X,M)$, because

$$\text{Cov}(Y,Z|X,M) = \text{Cov}(\alpha_D D + U, Z|X,M) = \text{Cov}(\alpha_D D, Z|X,M) = \text{Cov}(D,Z|X,M) \times \alpha_D$$

The second equality follows from U not being associated with Z conditional on both X and M when combining Assumption 1 with the conditional independence of W postulated in Assumption 2, which implies that M is not associated with U given X and thus exogenous. If the latter does not hold as in the right scenario of Figure 1, $\text{Cov}(U,Z|X,M)$ is different from zero because M is a function of both Z (via D) and (unobservables associated with) U given X. That is, $\text{Cov}(U,Z|X) = 0$ (which follows from Assumption 1) implies $\text{Cov}(U,Z|X,f(Z,X,U)) \neq 0$, where $f(Z,X,U)$ is a function of Z, X, U The partial effect $\alpha_D$is therefore not identified. Only in the case that $\beta_D = 0$ such that D does not affect M, it holds that $\text{Cov}(U,Z|X,M) = 0$ because M is not associated with Z given X. α_D, which in this case corresponds to the total effect of D on Y is therefore identified by an IV strategy conditional on X and M or conditional on X alone. The latter fact is easy to see from setting β_D to zero in equation (1).

Identifying the effect of education

Causal graph

This causal diagrams consist of a treatment variable D, which is now education, an outcome Y, and the unobserved terms U, V that effect treatment and outcome and may be arbitrarily associated with each other. As before, the instrument Z is confounded by pre-instrument covariates X. In contrast to the previous section, the instrument has now direct effects on two variables: First, it increases the likelihood to go to college due to draft avoiding behaviour. This effect will be used to identify the impact of education on labour market or health outcomes. The instrument also increases the likelihood to join the army, so that veteran status is now a post-instrument control variable M (Figure 3).

In this model, causal effects are only identified when controlling for both X and M: Since pre-instrument covariates are jointly associated with post-war outcomes and the unobservables, conditioning on X is required in an equivalent manner as in Figure 2. At the same time, the exclusion restriction of the instrument is not satisfied, which requires controlling for M.

Formal discussion based on a system of equations

We replace equation (2) in the model outlined in section 'Formal discussion based on a system of equations' by

$$M = \beta_z Z + \beta_X X + W, \tag{6}$$

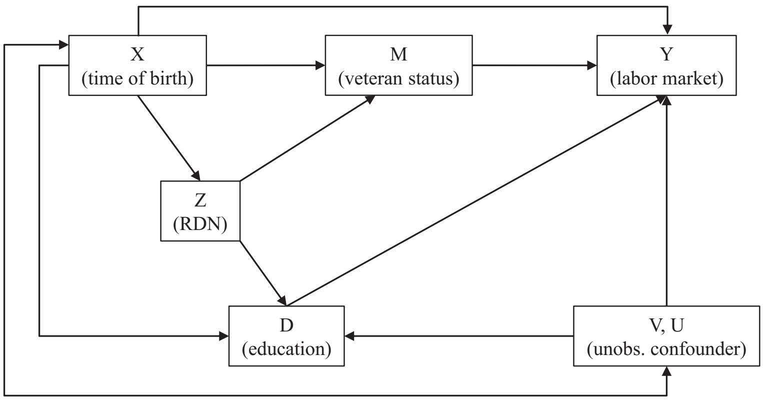


Figure 3 Causal diagram satisfying conditional IV independence conditional on (X, M)

such that the instrument now affects both D and M The effect of D on Y, α_D, cannot be identified by $\text{Cov}(Y, Z|X)/\text{Cov}(D, Z|X)$, because

$$\text{Cov}(Y, Z|X) = \text{Cov}(\alpha_D D + \alpha_M M + U, Z|X) = \text{Cov}(D, Z|X) \times \alpha_D + \text{Cov}(M, Z|X) \times \alpha_M$$

However, α_D is identified by

$$\text{Cov}(Y, Z|X, M) = \text{Cov}(\alpha_D D + U, Z|Xk, M) = \text{Cov}(\alpha_D D, Z|X, M) = \text{Cov}(D, Z|X, M) \times \alpha_D.$$

The second equality follows from the fact that U is not associated with Z conditional on both X and M, as implied by Assumptions 1 and 2.

Conditional IV independence is violated

Causal graph

The previous two sections demonstrated the identification of the causal effect of either veteran status or education in two separate cases, which was based on different sets of assumptions: section ‘Identifying the effect of veteran status when the instrument is confounded’ ruled out that the instrument Z had a direct effect on the post-instrument covariate M (in that case education) – otherwise the exclusion restriction would have been violated. This assumption was relaxed in section ‘Identifying the effect of education’, at the price of requiring the post-instrument covariate M (in that case veteran status) to be unconfounded by the unobserved variables U, V.

In this section, we combine features of the causal graphs of the two previous sections: We permit both variables of interest (veteran status and education) to be confounded by the unobserved variables U, V and the instrument to have a direct impact on both variables. We demonstrate that in this case, identification is not obtained under any of the discussed IV strategies. Because M is affected by Z, conditioning on M introduces and association between Z and V, U.[4] This entails collider/selection bias. However, not conditioning on

[4] As an example, suppose that the likelihood to join the army is negatively affected by the RDN and an individual’s aspirations to invest in formal education. Because the draft number is randomized, it should be independent of the aspirations. In the sample of veterans, however, individuals with high draft numbers (i.e. with no risk to be drafted) are more likely those who have low aspirations to invest in education. This spurious statistical association, however, is due to the selection process and does not reflect a statistical association in the general population.

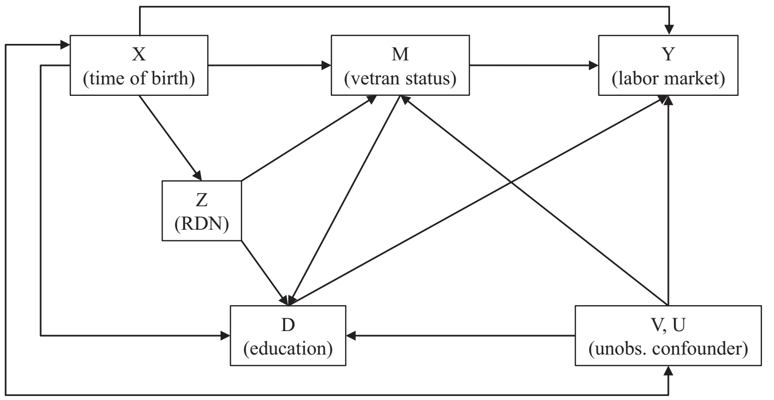


Figure 4 Causal diagram violating conditional IV independence

M violates the exclusion restriction, due to the causal path from the instrument to Y via M (Figure 4). It follows that neither the total, nor the direct effect are identified, no matter whether controlling for X alone, M alone, or both (X,M).

Unfortunately, such scenarios seem likely in many empirical problems where IV estimation is applied. With regard to the military draft, the educational deferment caused a strong incentive to be enrolled in college. Not controlling for education thus violates the exclusion restriction. However, education, labour market or health outcomes are the results of a common human capital investment process and are thus jointly determined (Fuchs, 1982). Since the underlying preference parameters are not fully observed in any conventional data source, education might be confounded by the same unobserved variables U, V. In such a scenario, a single instrument is thus not sufficient for identification.

Formal discussion based on a system of equations

Considering the model defined by equations (1), (3), (4), and (6) and assuming the satisfaction of Assumption 1, but the violation of Assumption 2, it follows that α_D, the effect of D on Y, cannot be identified. First, it does not correspond to $\mathrm{Cov}(Y,Z|X)/\mathrm{Cov}(D,Z|X)$, as outlined in section 'Formal discussion based on a system of equations'. Second, it cannot be obtained by $\mathrm{Cov}(Y,Z|X,M)/\mathrm{Cov}(D,Z|X,M)$, as

$$\mathrm{Cov}(Y,Z|X,M) = \mathrm{Cov}(\alpha_D D + U, Z|X,M) \neq \mathrm{Cov}(\alpha_D D, Z|X,M),$$

because $\mathrm{Cov}(U,Z|X,M) \neq 0$ under the violation of Assumption 2, see section 'Identifying the effect of veteran status when the instrument is confounded'.

IV. Empirical application

In the previous section, we discussed under which conditions draft risk can be either used as instrument for veteran status or education. We showed that none of these IV strategies can identify a causal effect if the following two conditions are violated: (a) the draft risk has no direct effect on post-instrument covariates and (b) the latter are not associated with the unobservables affecting the outcome. In this section, we provide empirical evidence which suggests that these conditions are indeed violated in our evaluation problem.

We use data from the 'Young Men in High School and Beyond' (YESB) survey (Bachman, 1999), which is a five-wave longitudinal study among a national sample of male students who were in 10th grade in fall 1966. This data set is particularly suited for three reasons: (i) More than 60% of the initial sample of 1966 were born in 1951 (1,380 respondents), the cohort exposed to the draft lottery of 1970. (ii) It is to the best of our knowledge the only publicly available data set that provides the exact birth date for at least a subsample of respondents, which is necessary to link draft lottery numbers (available from the Selective Service System: https://www.sss.gov/lotter1.htm) to individuals. (iii) In contrast to the previous literature that often relies on cross-sectional data, we observe respondents before the lottery took place. We are thus able to observe many pre-instrumental confounders that are typically unobserved in cross-sectional data.

Violation of condition (a): draft avoiding behaviour

The previous literature noticed stark education differences across individuals with high and low draft risks: Card and Lemieux (2001) for example found a strong correlation between the risk of induction faced by a cohort and the male-to-female ratio in college enrolment and completed schooling, which however flattens out for cohorts born after 1950. Angrist and Chen (2011) observed a similar pattern comparing men with low and high RDNs. The key problem in this literature is that the origin of the differences in education remains unclear: the gaps in schooling might either be explained by draft avoiding behaviour or be induced by the GI bill that finances continuous education after serving the military. While condition (a) would be violated in the former case, it would be satisfied in the latter.

Kuziemko (2010) is among the first providing convincing evidence for draft avoiding behaviour by focussing on the immediate behavioural change shortly after the lottery took place. She shows that white men born 1950 or earlier who randomly drew a low draft number in the 1969 lottery were more likely to attend college in Summer 1970 compared to white men with high draft number. The opposite is true for Afro-Americans or men from low-income groups. This supports the view that the lottery affects short term educational decisions which violate the exclusion restriction (i.e. condition a).[5]

Table 1 provides additional evidence for draft avoiding behaviour by focussing on completed education four years after the draft took place. We observe large and significant differences in completed education in 1974. Individuals with low draft numbers were 7 percentage points less likely to have at most a high school degree and 5 percentage points more likely to hold an associate's degree. Given that (i) veterans reported that they served for 36 months on average, (ii) no one in the sample started duty before summer 1970 (the large majority started military service in 1971), and (iii) an associate's degree takes 2 years of full time studying, the college degree effects of individuals with low draft numbers could not be induced by the GI bill. Therefore, they were most likely be driven by draft

[5]The analysis, however, is constrained by the fact that the draft number is only observed if the person was not in the army when the interview took place (about a half a year after the lottery). At the timing of the interview, many young men had already been drafted or enlisted due to low draft numbers. The draft numbers were thus not observed for the full population of interest and attrition was not random but associated with the key variable of interest. There are thus some concerns about potential sample selection bias.

TABLE 1

Completed education (1974)

	RDN > *125*	*RDN* ⩽ *125*	*Difference*
Total years of schooling	14.34	14.46	0.12
Highest degree: high school or less	0.56	0.49	−0.07**
Highest degree: associate's degree	0.10	0.15	0.05**
Highest degree: bachelor +	0.32	0.33	0.01

Notes: The sample includes individuals whose draft number is observed and who participated in the fifth wave (1974). Significance levels for the tests of equality in mean education outcomes: ***1%; **5%; *10%.

avoiding behaviour, which generally biases IV estimation strategies, using draft numbers as instrument for veteran status.

Violation of condition (b): education and veteran status are confounded

While the last section casts serious doubts on IV validity when estimating the effect of veteran status, we subsequently also challenge the IV-based estimation of education effects. The key problem is the requirement to condition on veteran status, which on the one hand controls for direct effects (i.e. not going through education) of the lottery on the outcome, but on the other hand most likely introduces collider bias due to the potential endogeneity of veteran status, i.e. its correlation with unobserved confounders affecting the outcome.

In contrast to other studies using the draft lottery instrument, our database contains some information on educational aspirations, which likely affect later life outcomes. We observe that individuals with low aspirations to invest in formal education were particularly likely to join the army in case of a low draft number. For instance, respondents reporting to have a college deferment in 1970 were only 12 percentage points more likely to report being a veteran in 1974, while individuals without a college deferment were 33 percentage points more likely.

Table 2 shows that once we control for veteran status, individuals with low draft numbers contain a higher share of individuals with high educational aspirations. It gives the results when regressing three variables measuring the willingness to invest in education prior to drafting on the instrument or on both the latter and veteran status. When using the instrument as the only regressor, coefficients are small and insignificant. When additionally controlling for veteran status in wave 5, the coefficients on the instrument are substantially larger and in two out of three cases significant at the 10% level. Conditional on veteran status, the instrument is therefore associated with (typically unobserved) educational aspirations, which likely determine outcomes like labour market success and health and thus violate condition (b).[6]

[6] Unfortunately, the analysis cannot be conducted for labour market outcomes measured prior to the lottery, as most respondents completed high school in 1969 and did not enter the labour market before the lottery took place.

TABLE 2

Suggestive evidence for collider bias

	Has college plans (W1)		*Student (W4)*		*College deferment (W4)*	
RDN < 125	0.04	0.07*	0.02	0.07*	0.00	0.06
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
Veteran (W5)		−0.14***		−0.20***		−0.26***
		(0.05)		(0.05)		(0.06)
Constant	0.68***	0.69***	0.67***	0.68***	0.54***	0.55***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
No. observations	865		865		830	
Equality of coefficient	6.49**		11.07***		15.83***	

Notes: OLS coefficients on the draft number dummy, veteran status, and a constant when using various background variables as outcomes (whether or not the respondent had plans to enroll in college in 1966, if the person was already student at the time of the interview in 1970, and if the person had a college deferment at the time of the interview in 1970). The sample includes individuals whose draft number is observed and who participated in the fifth wave (1974). Further control variables are veteran status, month of birth and race. Standard errors are in parentheses. Significance levels: ***1%; **5%; *10%.

V. Conclusion

This paper analysed different conditional instrumental variable (IV) strategies controlling for observed covariates to understand under which assumptions in terms of the behaviour of pre- and post-instrument control variables identification is (not) obtained. Throughout this paper, we used the Vietnam draft as running example for an IV approach for evaluating the effects of military service or education on later life outcomes such as labour market performance and health. Furthermore, in an empirical application to the ‘Young Men in High School and Beyond’ survey, we demonstrated that commonly used conditional IV strategies to estimate the impact of draft induced military service or education may be severely biased by draft avoiding behaviour (which violates the exclusion restriction) and collider/selection bias that arises if one controls for a post-instrument variable like veteran status. These results suggest that researchers need to carefully think about possible causal relations of all variables entering the IV model and the role of (the timing of) control variables for obtaining identification. Causal diagrams as used in this paper can be a powerful tool for visualizing such issues in complex causal models.

Final Manuscript Received: May 2016

Appendix

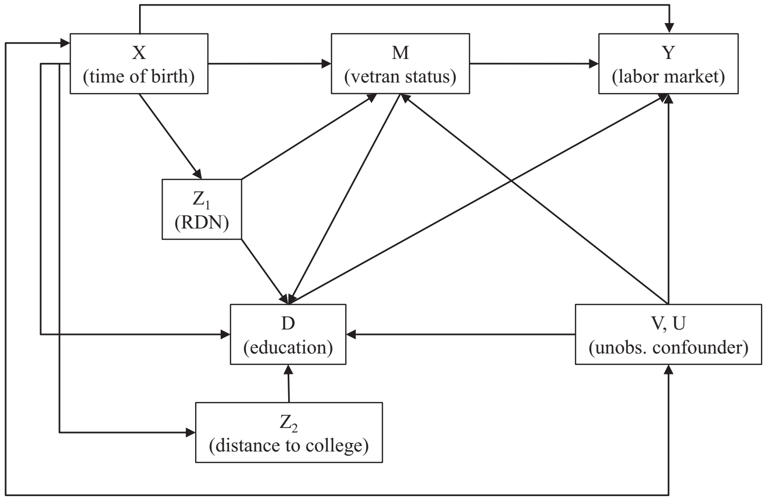


Figure A1 Causal diagram compatible with Chaudhuri and Rose (2009)

References

Angrist, J. D. (1990). 'Lifetime earnings and the Vietnam Era draft lottery: evidence from social security administrative records', *American Economic Review*, Vol. 80, pp. 313–336.

Angrist, J. D. (1991). 'The draft lottery and voluntary enlistment in the Vietnam Era', *Journal of the American Statistical Association*, Vol. 86, pp. 584–595.

Angrist, J. D. and Chen, S. H. (2011). 'Schooling and the Vietnam–Era GI bill: evidence from the draft lottery', *American Economic Journal Applied Economics*, Vol. 3, pp. 96–118.

Angrist, J. D. and Krueger, A. B. (1992). *Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery*, National Bureau of Economic Research, Cambridge, MA.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricists Companion*, Princeton University Press, Princeton, NJ.

Angrist, J. D., Chen, S. H. and Frandsen, B. R. (2010). 'Did Vietnam veterans get sicker in the 1990s? The complicated effects of military service on self-reported health', *Journal of Public Economics*, Vol. 94, pp. 824–837.

Bachman, J. G. (1999). *Young Men in High School and Beyond: A Summary of Findings from the Youth in Transition Project, 1966–74*, Inter-University Consortium for Political and Social Research, Ann Arbor, MI.

Bazzi, S. and Clemens, M. A. (2013). 'Blunt instruments: avoiding common pitfalls in identifying the causes of economic growth', *American Economic Journal: Macroeconomics*, Vol. 5, pp. 152–186.

Bedard, K. and Deschênes, O. (2006). 'Association the long-term impact of military service on health: evidence from World War II and Korean War Veterans', *American Economic Review*, Vol. 96, pp. 176–194.

Buckles, K. S. and Hungerman, D. M. (2013). 'Season of birth and later outcomes: old questions, new answers', *Review of Economics and Statistics*, Vol. 95, pp. 711–724.

Card, D. and Lemieux, T. (2001). 'Going to college to avoid the draft: the unintended legacy of the Vietnam War', *American Economic Review*, Vol. 91, pp. 97–102.

Chalak, K. and White, H. (2011). 'Viewpoint: an extended class of instrumental variables for the estimation of causal effects', *Canadian Journal of Economics*, Vol. 44, pp. 1–51.

Chaudhuri, S. and Rose, E. (2009). *Estimating the Veteran Effect with Endogenous Schooling When Instruments Are Potentially Weak*. IZA Discussion Paper No. 4203.

Conley, D. and Heerwig, J. (2011). *The War at Home: Effects of Vietnam–Era Military Service on Post–War Household Stability*. National Bureau of Economic Research, Cambridge, MA.

Dobkin, C. and Shabani, R. (2009). 'The health effects of military service: evidence from the Vietnam draft', *Economic Inquiry*, Vol. 47, pp. 69–80.

Elwert, F. and Winship, C. (2014). 'Endogenous selection bias: the problem of conditioning on a collider variable', *Annual Review of Sociology*, Vol. 40, pp. 31–53.

Fienberg, S. (1971). 'Randomization and social affairs: the 1970 draft lottery', *Science*, Vol. 171, pp. 255–261.

Fuchs, V. R. (1982). 'Time preference and health: an exploratory study', in Fuchs V. R. (ed), *Economic Aspects of Health*, Chicago, IL: University of Chicago Press, pp. 93–120.

Gimbel, C. and Booth, A. (1996). 'Who fought in Vietnam?', *Social Forces*, Vol. 74, pp. 1137–1157.

Grimard, F. and Parent, D. (2007). 'Education and smoking: were vietnam war draft avoiders also more likely to avoid smoking?', *Journal of Health Economics*, Vol. 26, pp. 896–926.

Heckman, J. J. (1979). 'Sample selection bias as a specification error', *Econometrica*, Vol. 47, pp. 153–161.

Huber, M. (2014). 'Identifying causal mechanisms (primarily) based on inverse probability weighting', *Journal of Applied Econometrics*, Vol. 29, pp. 920–943.

Kuziemko, I. (2010). *Did the Vietnam Draft Increase Human Capital Dispersion? Draft–Avoidance Behavior by Race and Class*. Princeton and NBER Working Paper.

Lechner, M. (2008). 'A note on endogenous control variables in causal studies', *Statistics and Probability Letters*, Vol. 78, pp. 190–195.

Lindo, J. M. and Stoecker, C. (2014). 'Drawn into violence: evidence on "what makes a criminal" from the Vietnam draft lotteries', *Economic Inquir*, Vol. 52, pp. 239–258.

de Luna, X. and Johansson, P. (2014). 'Testing for the unconfoundedness assumption using an instrumental assumption', *Journal of Causal Inference*, . Vol. 2, pp. 187–199.

Lundberg, S. (2005). 'Sons, daughters, and parental behaviour', *Oxford Review of Economic Policy*, Vol. 21, pp. 340–356.

Malamud, O. and Wozniak, A. K. (2010). *The Impact of College Education on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk*, National Bureau of Economic Research, Cambridge, MA.

Pearl, J. (1995). 'Causal diagrams for empirical research', *Biometrika*, Vol. 4, pp. 669–688.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge.

Rosenbaum, P. R. (1984). 'The consequences of adjustment for a concomitant variable that has been affected by the treatment', *Journal of the Royal Statistical Society, A*, Vol. 147, pp. 656–666.

de Walque, D. (2007). 'Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education', *Journal of Health Economics*, Vol. 26, pp. 877–895.

White, H. and Lu, X. (2011). 'Causal diagrams for treatment effect estimation with application to efficient covariate selection', *Review of Economics and Statistics*, Vol. 93, pp. 1453–1459.