Research Article

Olena Y. Nizalova* and Irina Murtazashvili

Exogenous Treatment and Endogenous Factors: Vanishing of Omitted Variable Bias on the Interaction Term

Abstract: Whether interested in the differential impact of a particular factor in various institutional settings or in the heterogeneous effect of policy or random experiment, the empirical researcher confronts a problem if the factor of interest is correlated with an omitted variable. This paper presents the circumstances under which it is possible to arrive at a consistent estimate of the mentioned effect. We find that if the source of heterogeneity and omitted variable are jointly independent of policy or treatment, then the OLS estimate on the interaction term between the treatment and endogenous factor turns out to be consistent.

Keywords: heterogeneity; interaction term; omitted variable bias; random experiments; treatment effect.

DOI 10.1515/jem-2013-0012
Previously published online September 19, 2014

1 Introduction

Significant increase in the use of random experiments in the development economics and natural experiments throughout other fields of economics is raising the question of whether it is possible to obtain a consistent estimate of the heterogeneous treatment effect if the heterogeneity is occurring along the lines of a factor which is correlated with some omitted variable(s). Likewise, empirical researchers are often interested in estimation of the differential impact of a particular factor (which maybe correlated with omitted variables) in various institutional settings. These two situations are similar if the policy variable or assignment to the treatment group is uncorrelated with either the factor of interest (source of heterogeneity) or with the omitted variable inasmuch as the goal is to estimate the coefficient on the interaction term between the policy/treatment variable and the factor of interest which is correlated with the error term.

The textbook approach to econometric modeling suggests that we ought to include all the relevant variables into a model. The justification of this approach is due to possible (partial) correlations among the explanatory variables. Indeed, every standard econometric textbook shows, if included regressors are partially correlated with an excluded additional explanatory variable, the exclusion of this additional relevant regressor will result in omitted variable bias.[1]

This straightforward theoretical result is of serious consequence for data analysts, since applied researchers are rarely able to follow the suggestion to include all the relevant explanatory variable. In reality, we

1 It is worth reminding another relevant standard textbook fact: excluding an explanatory variable that is partially uncorrelated with included regressors has no effect on unbiasedness and consistency of the OLS estimates.

***Corresponding author: Olena Y. Nizalova,** PSSRU/CHSS, University of Kent, George Allen Wing, Cornwallis Building, Canterbury, Kent, Ct2 7NF, Phone: +44 1227 824966, E-mail: O.Nizalova@kent.ac.uk
Irina Murtazashvili: School of Economics, LeBow College of Business, Drexel University, Philadelphia, PA 19104, USA

cannot always include all the omitted variables for various reasons, often due to their unobservability. Unless there is an instrumental variable (IV) available, there is little hope to get consistent estimates of the model parameters then.

As an alternative to the IV approach, one can assess the magnitude or at least the direction of the bias. However, theoretical textbooks' discussions about omitted variable bias always focus on the example when the true model contains two variables (in addition to the constant term), but the estimated model omits one variable, which is correlated with the regressor of interest. This setup allows researchers to talk about the direction of the bias and speculate whether the biased OLS estimate helps in understanding the issue at hand or one should definitely be searching for a way to obtain more consistent estimates. But every textbook consideration of the issue concludes with the warning that in the case of three or more variables in the model, it is difficult to tell what would be the direction of the bias. This applies to the estimation of the heterogeneous treatment effect since there are at least four variables in this setting (in addition to the constant term): an endogenous factor, an omitted variable correlated with this endogenous factor, an exogenous treatment,[2] and an interaction term between the treatment and endogenous factor.

A natural question that comes to mind in this case is whether there are at least some situations when the exclusion of the relevant variable is of not such a severe consequence. Is there a scenario under which the unobserved covariate correlated with the included regressors does not cause much trouble (at least) for some of the model parameters that are of interest? It turns out that this situation is indeed possible and quite common in applied works. Let all the regressors but the exogenous regressor of main interest and the interaction term between this exogenous regressor and an endogenous covariate to be jointly independent of the exogenous regressor of the main interest.[3] Then, the OLS estimate of the coefficient on this interaction term is consistent. Therefore, one can use this result to inform policy makers of the differential impact of some endogenous factors in different policy settings, or about heterogeneous treatment effect when the source of heterogeneity is endogenous, provided that the endogenous factor of interest and the unobservable are jointly independent of the policy/treatment.

To the best of our knowledge, while not necessarily surprising to theoretical econometricians, consistency of the OLS estimate of the coefficient for the interaction between a policy/treatment variable and an observed endogenous factor when the covariate and the unobservable are jointly independent of the policy/treatment has not been emphasized previously. Here we derive this rather important result that is particularly relevant for practitioners explicitly. The rest of the paper is structured in the following way. Section 2 describes the relevant applications. Section 3 provides the econometric result. Using Monte Carlo simulations, Section 4 illustrates the finite sample properties of the OLS estimator in our setting. Conclusions follow in Section 5.

2 Some Examples of Relevant Applications

Earlier works which evaluated the effects of large scale random experiments and those which exploited the so-called natural experiments mostly focused on the estimation of the treatment effect only. One of the exceptions we found dates back to 1991 and describes the experimental evidence on the effects of double-blind versus single-blind reviewing on the probability of acceptance of a paper for publication in the American Economic Review (Blank 1991).

The AER experiment was held over the period 1987–1989 and resulted into a sample of 1498 papers with completed referee reports, which were either double-blind or single-blind through a random assignment. The results suggested that the double-blind procedure is stricter, which is confirmed by a significantly lower acceptance rate and more critical referee reports. However, the emphasis of the paper is not on the overall

2 We call treatment exogenous as we assume that the source of heterogeneity and omitted variable(s) are jointly independent of the treatment.

3 We also discuss a weaker set of conditions later in the paper.

effect of the double-blind refereeing, but rather on the heterogeneous impact of the treatment, which is the focus of this paper. In particular, some earlier studies found that women have higher acceptance rates in double-blind journals (Ferber and Teiman 1980), and this was chosen as one of the important dimensions of heterogeneity. Other dimensions included the rank of the university and indicators whether the institution is US nonacademic or foreign. Clearly, gender is likely to be correlated with other important factors, which were not observed in the experiment, such as age and experience in the profession. Likewise, being in a higher ranked university maybe the result of the overall higher unobserved productivity.

A simplified relation between the acceptance rates and assignment to the review group studied by Blank (1991) can be described as:

$$y=\beta_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\varepsilon^*, \tag{1}$$

where $x_2=x_3\cdot x_4$ is the interaction term, and $\varepsilon^*=\varepsilon+c$. Here, x_3 specifies the university rank,[4] x_4 is an indicator of the double-blind treatment, c is the unobserved individual-specific effect, and ε is the idiosyncratic error. For simplicity of illustration, we specify a model with only three explanatory variables, while Blank (1991) estimates a more complex model that includes several interaction terms. We generalize our discussion of model (1) in the next section.

The coefficients on interaction terms studied in the AER experiment by Blank (1991) turned out to be statistically insignificant, suggesting no benefits of double-blind refereeing to either women or authors from lower-ranked universities. But can this finding be trusted? The author states that the coefficients on the interaction terms "should be robust to the inclusion of any other variables in the model, since they come from two experimental samples that are identical in all other characteristics" (Blank 1991, 1054). At the same time with respect to the main effects of gender and the university rank, the author claims that "it is not clear how to interpret the coefficients on these variables, because they are contaminated by excluded variables" (Blank 1991, 1055). These statements are indications of what we are to prove explicitly in this paper: the consistency of the estimates of the heterogeneous impact of random treatment/exogenous policy when the heterogeneity occurs along the lines of a factor correlated with the omitted variable(s).

In recent years a considerable number of works has appeared which either directly investigate the heterogeneity of treatment effect or point to the possibility of its existence. However, the studies which do estimate the heterogeneous effects are more reserved than Blank (1991) with respect to the discussion of the consistency of the estimates.

Blau et al. (2010) report on the impact of a trial in which the Committee on the Status of Women in the Economics Profession (CSWEP) randomly chose the participants of the CSWEP Mentoring Program (CeMENT) which "aimed at assisting female junior faculty in preparing themselves for the tenure hurdle." The authors find that in 3–5 years after the Program participants have higher likelihood of having any top-tier publication and more publications in general, as well as more federal grants. As the rate of acceptance to the journals may depend on the rank of the university (Blank 1991), it may be interesting to investigate whether the impact of the CeMENT is different for junior female faculty from low-rank versus high rank universities.

A recent study by Glewwe, Kremer, and Moulin (2009) focuses on the evaluation of a randomized trial in rural Kenya estimating the effect of provision of free textbooks on the students' test scores. Compared to the earlier literature on the effect of the textbook provision on the test scores, the authors find no significant treatment effect. However, when taking into account the heterogeneity by the past test scores, they reach the conclusion that the best students do benefit from the textbook provision. The study has a cross-sectional set-up and therefore the authors could not control for students' ability. The previous test scores, likewise the current test scores, are clearly correlated with the unobserved ability. Therefore, the authors study the heterogeneity of the treatment effect along the lines of a factor which is correlated with the error term. Similarly, Banerjee et al. (2007) evaluate the two randomized experiments in India where a remedial education program hired young women to teach students lagging behind in basic literacy and numeracy skills. They

4 Although the university rank is represented by a set of indicators in (Blank 1991), we use one variable, x_3.

also consider the previous test scores as the source of the heterogeneity of impact by dividing the sample into terciles according to the past score distribution. The largest gains are experienced by children at the bottom of the test-score distribution.

Banerjee et al. (2014) estimate the impact of a randomized introduction of microcredit in a new market. They find that households with an existing business at the time of the program invest more in durable goods. Moreover, households with high propensity to become business owners see a decrease in nondurable consumption, while households with low propensity to become business owners show an increase in nondurable spending. The study is again set up as a cross-section and there is a considerable room for omitting variables which determine past business ownership and current propensity to become a business owner and the consumption patterns. People who are already business owners or have a higher potential to become ones are potentially different from the rest of the population in characteristics which may as well determine the spending patterns.

3 Econometric Result

In practice, we mostly encounter regression equations that include more than three explanatory variables (in addition to the constant term). Therefore, we re-write equation (1) in more general terms, and proceed further in the context of the AER experiment in Blank (1991). Following a standard approach of reparameterizing models with interactions, we demean vectors of endogenous and treatment covariates in the interaction terms to get:

$$y_i = \alpha_1 + \ddot{\mathbf{x}}_{2i}\boldsymbol{\beta}_2 + \mathbf{x}_{3i}\boldsymbol{\alpha}_3 + \mathbf{x}_{4i}\boldsymbol{\alpha}_4 + \varepsilon_i^*, \quad (2)$$

where $\varepsilon^* = \varepsilon + c$, c is some unobserved heterogeneity, $\mathbf{x}_3$ is a vector containing the endogenous covariates correlated with c, $\mathbf{x}_4$ is a vector of the treatment variables, and $\ddot{\mathbf{x}}_2 = ((\mathbf{x}_{3i} - \boldsymbol{\mu}_3) \otimes (\mathbf{x}_{4i} - \boldsymbol{\mu}_4))$ is a vector of interaction terms that were constructed using demeaned $\mathbf{x}_3$ and $\mathbf{x}_4$, where $\boldsymbol{\mu}_j = E(\mathbf{x}_{ji})$, $j=3, 4$, and $\otimes$ is the Kronecker product. Recall that reparametarization is used in models with interactions to ease the interpretation of the coefficients on the individual covariates, $\mathbf{x}_3$ and $\mathbf{x}_4$. Note that while demeaning affects the coefficients on $\mathbf{x}_3$ and $\mathbf{x}_4$ in the original equation without reparametarization, it does not affect the coefficients of the interaction terms – $\boldsymbol{\beta}_2$ – the parameters of our interest.[5] In the context of the AER experiment in Blank (1991), $\mathbf{x}_3$ contains a set of indicators for the university rank, and indicators for the nature of the institutions and female authors, $\mathbf{x}_4$ is a set of the dummies for the double-blind refereeing, and c represents age, experience in the profession, and author productivity unobserved in the experiment.

The question that Blank (1991) raises is whether the double-blind reviewing affects the acceptance rates differently depending on the university rank, author gender, and the nature of the institution, i.e., whether $\boldsymbol{\beta}_2$ is statistically different from zero. However, the unobserved personality traits are correlated with the university rank, gender and the nature of the institution. Standard econometric wisdom suggests that in a cross-sectional setting the estimates of all the parameters will be inconsistent since Corr($\mathbf{x}_3$, ε^*)$\neq\mathbf{0}$. But is this indeed the case?

We exploit general formula (9) provided in the Appendix to get the probability limit of $\hat{\beta}_2$ in equation (2). Then,

$$\begin{aligned}\text{plim}(\hat{\beta}_2) &= \hat{\beta}_2 + [\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\tilde{\ddot{\mathbf{X}}}_2)]^{-1}\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'(\tilde{\mathbf{X}}_3, \tilde{\mathbf{X}}_4)) + [\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\tilde{\ddot{\mathbf{X}}}_2)]^{-1}\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\widetilde{\boldsymbol{\varepsilon}^*}) \\ &= \boldsymbol{\beta}_2 + [\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\tilde{\ddot{\mathbf{X}}}_2)]^{-1}\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\tilde{\mathbf{X}}_3, \tilde{\ddot{\mathbf{X}}}_2'\tilde{\mathbf{X}}_4) + [\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\tilde{\ddot{\mathbf{X}}}_2)]^{-1}\,\text{plim}(\tilde{\ddot{\mathbf{X}}}_2'\widetilde{\boldsymbol{\varepsilon}^*}),\end{aligned} \quad (3)$$

where we use $\tilde{\mathbf{Z}}$ to denote demeaned matrix $\mathbf{Z}$ with all observations stacked together for any $\mathbf{Z}$.

5 The exact relations between α_j and β_j, where j=1, 3, 4, are easily derivable.

We are interested in cases when we can assume independence between $\mathbf{x}_4$ and $(\mathbf{x}_3, \varepsilon^*)$. The independence between $\mathbf{x}_4$ and $(\mathbf{x}_3, \varepsilon^*)$ allows us to obtain the following expressions:

$$E(\tilde{\ddot{\mathbf{x}}}_{2i}'\tilde{\varepsilon}_i^*)=E[[((\mathbf{x}_{3i}-\boldsymbol{\mu}_3)\otimes(\mathbf{x}_{4i}-\boldsymbol{\mu}_4))-\boldsymbol{\mu}_{34}]'\tilde{\varepsilon}_i^*]=E[(\mathbf{x}_{3i}-\boldsymbol{\mu}_3)'\tilde{\varepsilon}_i^*]\otimes E(\mathbf{x}_{4i}-\boldsymbol{\mu}_4)'-\boldsymbol{\mu}_{34}'E(\tilde{\varepsilon}_i^*)=\mathbf{0}, \quad (4)$$

and

$$E(\tilde{\ddot{\mathbf{x}}}_{2i}'\tilde{\mathbf{x}}_{ji})=E[[(\mathbf{x}_{3i}-\boldsymbol{\mu}_3)\otimes(\mathbf{x}_{4i}-\boldsymbol{\mu}_4)]'(\mathbf{x}_{ji}-\boldsymbol{\mu}_j)]-\boldsymbol{\mu}_{34}'E(\mathbf{x}_{ji}-\boldsymbol{\mu}_j)=[E(\mathbf{x}_{3i}-\boldsymbol{\mu}_3)'\otimes E(\mathbf{x}_{4i}-\boldsymbol{\mu}_4)']E(\mathbf{x}_{ji}-\boldsymbol{\mu}_j)=\mathbf{0}, j=3,4, \quad (5)$$

where $\mu_{34}=E[(\mathbf{x}_3-\boldsymbol{\mu}_3)\otimes(\mathbf{x}_4-\boldsymbol{\mu}_4)]$. The Law of Iterated Expectations and the assumption of independence between $\mathbf{x}_4$ and $(\mathbf{x}_3, \varepsilon^*)$ are utilized to obtain the second equalities in both (4) and (5). We also use a standard assumption that $E(\varepsilon^*)=0$ when deriving (4). Applying (4) and (5) to equation (3), we conclude that $\text{plim}(\hat{\beta}_2)=\beta_2$ when $\mathbf{x}_4$ is independent from $(\mathbf{x}_3, \varepsilon^*)$. Therefore, the OLS estimates of the coefficients on the interaction terms in equation (2) are consistent when $\mathbf{x}_4$ is independent from $(\mathbf{x}_3, \varepsilon^*)$. Additionally, note that we can follow the same logic with general formula (9) (in the Appendix) to prove the consistency of the OLS estimates of the main treatment effects.

We have showed that the OLS coefficient estimates of $(\mathbf{x}_3\otimes\mathbf{x}_4)$ and $\mathbf{x}_4$ in equation (2) are consistent under independence of $\mathbf{x}_4$ and $(\mathbf{x}_3, \varepsilon^*)$, which is actually stronger than necessary to guarantee this result. For consistency, it would be sufficient to have either $f(\mathbf{x}_3|\mathbf{x}_4, \varepsilon^*)=f(\mathbf{x}_3|\varepsilon^*)$ or $f(\varepsilon^*|\mathbf{x}_3, \mathbf{x}_4)=f(\varepsilon^*|\mathbf{x}_3)$ in combination with $\mathbf{x}_4$ being independent of either ε^* or $\mathbf{x}_3$, respectively.

Let us revisit the Blank (1991) study. The question of interest there is estimating the differences in the effect of the double-blind reviewing procedure for different groups of researchers. The author is after the coefficient estimates of the interaction terms between the university rank, author gender, indicators whether the institution is US nonacademic or foreign, and the variable identifying the sample randomly assigned to the double-blind reviewing. While there are valid reasons to suspect that the university rank (or author gender and nature of the institution) is correlated with the unobservables (say, productivity of the author), this treatment is independent of the university rank, author gender, indicators for whether the institution is US nonacademic or foreign as well as productivity of the authors. These independences guarantee that the OLS estimates of the interaction terms between university rank, gender, indicators for the nature of the institutions, and treatment dummies are consistent as we show above.

4 Small Sample Behavior of the OLS Estimator in Our Setting

In this section we employ Monte Carlo simulations to draw the data and check the finite sample properties of the OLS estimator under the assumptions of our interest. We use 1000 replications to study this question when two sample sizes: N=100 and N=1000. The data generating process (DGP) employed is:

$$y_i=1+2(r_i-\bar{r})\cdot(d_i-\bar{d})+3r_i+4d_i+5f_i+6s_i+7n_i+8c_i+u_i, \quad (6)$$

where $\bar{r}$ and $\bar{d}$ are sample means of r_i and d_i, respectively. Here, r_i (university rank) and u_i (idiosyncratic error) are generated as independent Normal (0, 1). The unobserved heterogeneity, c_i, is generated as $c_i=0.5r_i+e_i^c$, where $e_i^c\sim$Normal (0, 1). The exogenous treatment, d_i, is generated as Bernouli (0.5).

We consider three possibilities for additional regressors: (1) regressors independent of c_i (e.g. gender of the referee), (2) regressors correlated with d_i but uncorrelated with c_i,[6] (3) regressors with non-zero simple correlation with c_i (e.g. gender of the author). Case (1) is represented by $f_i\sim$Bernoulli (0.5). Case (2) is represented by $s_i=0.5d_i-1+e_i^s$, where $e_i^s\sim\text{Discrete uniform}(0,3)$. For case (3), we consider two DGPs for n_i – the rank of

6 It is difficult to think of such a regressor in the AER experiment, but generally it is possible to have such variables.

Table 1: OLS Estimation Results for $(\beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7)'=(2, 3, 4, 5, 6, 7)'$.

# of Regressors:		**6**	**7**	**6**	**7**	**6**	**7**	**6**	**7**
		***N*=100**		***N*=1000**		***N*=100**		***N*=1000**	
		(A): $n_i=0.5r_i+e_i^n$				**(B): $n_i=0.5c_i+e_i^n$**			
(1)	$\hat{\beta}_2$	2.071	2.009	1.989	1.999	2.080	2.009	1.987	1.999
	$SE(\hat{\beta}_2)$	(1.677)	(0.211)	(0.512)	(0.064)	(1.506)	(0.211)	(0.459)	(0.064)
(2)	$\hat{\beta}_3$	6.991	3.001	7.012	2.999	6.215	3.001	6.198	2.999
	$SE(\hat{\beta}_3)$	(0.934)	(0.129)	(0.287)	(0.039)	(0.771)	(0.118)	(0.235)	(0.036)
(3)	$\hat{\beta}_4$	3.913	4.006	3.972	4.003	3.929	4.006	3.983	4.003
	$SE(\hat{\beta}_4)$	(1.670)	(0.210)	(0.524)	(0.065)	(1.500)	(0.210)	(0.470)	(0.065)
(4)	$\hat{\beta}_5$	4.986	5.001	4.992	5.000	5.014	5.001	4.990	5.000
	$SE(\hat{\beta}_5)$	(1.638)	(0.206)	(0.512)	(0.063)	(1.471)	(0.206)	(0.458)	(0.063)
(5)	$\hat{\beta}_6$	6.054	5.998	5.997	6.000	6.030	5.998	5.996	6.000
	$SE(\hat{\beta}_6)$	(0.733)	(0.092)	(0.256)	(0.028)	(0.658)	(0.092)	(0.205)	(0.028)
(6)	$\hat{\beta}_7$	7.006	7.000	6.995	7.000	10.139	7.000	10.209	7.000
	$SE(\hat{\beta}_7)$	(0.826)	(0.104)	(0.256)	(0.032)	(0.665)	(0.104)	(0.205)	(0.032)
(7)	$RMSE(\hat{\beta}_2)$	1.698	0.207	0.498	0.065	1.536	0.207	0.443	0.065
(8)	$SD(\hat{\beta}_2)$	1.698	0.207	0.498	0.065	1.535	0.207	0.443	0.065
(9)	$LQ(\hat{\beta}_2)$	0.915	1.870	1.648	1.951	0.959	1.870	1.691	1.951
(10)	$Median(\hat{\beta}_2)$	2.071	1.989	1.985	1.998	2.097	1.989	1.972	1.998
(11)	$UQ(\hat{\beta}_2)$	3.244	2.145	2.340	2.046	3.151	2.145	2.287	2.046

Notes: Odd columns report results for the estimating equation with six regressors, while even columns – for the estimating equation with all seven regressors. Rows (1) through (6) contain means of OLS slope estimates and their corresponding standard errors from 1000 replications. Rows (7) through (11) contain the root mean squared error (RMSE), standard deviation (SD), lower quartile (LQ), median, and upper quartile (UQ) for $\hat{\beta}_2$ – our main coefficient of interest – from 1000 replications. Also, the first four columns report the results when n_i is generated according to DGP (A), while the last four columns – according to DGP (B).

the school granting doctorate to the author:[7] (A) $n_i=0.5r_i+e_i^n$, and (B) $n_i=0.5c_i+e_i^n$, where $e_i^s \sim Normal(0, 1)$. These two DGPs result in non-zero simple correlation between c_i and n_i. However, the partial correlation between n_i and c_i, i.e., correlation net of the effect of the other included regressors (in particular, r_i), is zero for DGP (A), while it is clearly not for DGP (B).

Table 1 presents simulation results. We consider two regressions: with six (c_i is excluded) and seven (c_i is included) regressors (in addition to the constant term). Note that $\hat{\beta}_3$ from the model with six regressors is inconsistent regardless of N and DGP for n_i. The fact that $Corr(s_i, d_i)\neq 0$ has no effect on any of the OLS estimates in all cases, since these variables are independent of c_i. Similarly, $\hat{\beta}_5$ is always consistent.

Clearly, when seven regressors are included all estimates are consistent. More importantly, when only six regressors are used, $\hat{\beta}_2$ and $\hat{\beta}_4$ are consistent and essentially unbiased,[8] while the consistency (and the extent of bias) of $\hat{\beta}_3$ and $\hat{\beta}_7$ depends on the (partial) correlations $Corr(r_i, c_i)$ and $Corr(n_i, c_i)$, respectively. The simulation findings are unambiguous: when the partial correlation between the unobserved heterogeneity and some included regressor is different from zero, the OLS slope estimate of that included regressor is the only estimate which is inconsistent, and its bias does not disappear as $N\rightarrow\infty$.

5 Conclusions

Increasing interest in the heterogeneity of the impact in policy evaluation and random experiment settings leads to a question of whether the estimates are consistent when the source of heterogeneity is correlated

7 In case of multiple authors, this can be measured by the highest rank of the schools granting doctorate among all co-authors.
8 We report the detailed results for $\hat{\beta}_2$ only but the results for $\hat{\beta}_4$ are available upon request.

with some omitted variable(s). This paper presents the conditions under which it is possible to arrive at a consistent OLS estimate of the mentioned effect. We explicitly show that if the source(s) of heterogeneity and omitted variable(s) are jointly independent of the policy/treatment(s), then the OLS estimates of the main treatment effect(s) and the coefficient(s) on the interaction term(s) between the treatment(s) and endogenous factor(s) are still consistent. This matter has not been emphasized explicitly before, yet represents a significant interest for applied and policy research circles. We discuss the relevant applications and provide simulation evidence for the finite sample properties of the OLS estimator in such a setting.

Acknowledgements: This paper has benefited from valuable comments and suggestions of Jeffrey Wooldridge. We are also thankful to Tom Coupe, Soiliou Namoro, Jean-Francois Richard, Peter Schmidt for helpful comments and discussions. The authors acknowledge financial support from a UCSUR Steven Manners Research Development Award (University of Pittsburgh).

Appendix

The popular econometric textbook by Greene (2007) derives the following general result. Suppose the correct specification of the regression model for all observations stacked together is

$$\mathbf{y}=\mathbf{i}\gamma_1+\mathbf{V}\gamma_2+\mathbf{W}\gamma_3+\varepsilon^*, \tag{7}$$

where $\mathbf{i}$ is a $n\times 1$ vector of ones. Premultiplying equation (7) by matrix $\mathbf{M}=\mathbf{I}-\mathbf{i}(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'$, where $\mathbf{I}$ is an $n\times n$ identity matrix, yields a demeaned version of the original model:

$$\tilde{\mathbf{y}}=\tilde{\mathbf{V}}\gamma_2+\tilde{\mathbf{W}}\gamma_3+\tilde{\varepsilon}^*, \tag{8}$$

where $\tilde{\mathbf{Z}}$ denotes mean-differenced $\mathbf{Z}$ for any $\mathbf{Z}$.[9] Further, suppose we do not include $\mathbf{W}$ into our regression (7) and, therefore, estimate $\tilde{\mathbf{y}}=\tilde{\mathbf{V}}\gamma_2+\mathbf{u}$, where $\mathbf{u}=\widetilde{\tilde{\mathbf{W}}\gamma_3+\varepsilon^*}$. We make a standard assumption that $E(\varepsilon^*)=0$. Then, we can modify the omitted variable formula from Greene (2007) to report the probability limit of $\tilde{\gamma}_2$:

$$\text{plim}(\hat{\gamma}_2)=\gamma_2+\text{plim}(\tilde{\mathbf{V}}'\tilde{\mathbf{V}})^{-1}\tilde{\mathbf{V}}'\tilde{\mathbf{W}}\cdot\gamma_3+\text{plim}(\tilde{\mathbf{V}}'\tilde{\mathbf{V}})^{-1}\tilde{\mathbf{V}}'\tilde{\varepsilon}^*. \tag{9}$$

References

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2014. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." Retrieved September 1, 2014. http://econ-www.mit.edu/files/5993.
Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics* 122 (3), 1235–1264.
Blank, Rebecca M. 1991. "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence From The American Economic Review." *American Economic Review* 81 (5): 1041–1067.
Blau, Francine D., Janet M. Currie, Rachel T. A. Croson, and Donna K. Ginther. 2010. "Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial." *American Economic Review* 100 (2): 348–352.
Ferber, Marianne A., and Michelle Teiman. 1980. "Are Women Economists at a Disadvantage in Publishing Journal Articles?" *Eastern Economic Journal* 6(3–4): 189–194.
Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* 1 (1): 112–135.
Greene, William H. 2007. *Econometric Analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

9 Note that we are not able to directly estimate the intercept γ_1 from the mean-differenced model.