Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: https://www.tandfonline.com/loi/lecr20

OLS and IV estimation of regression models including endogenous interaction terms

Maurice J. G. Bun & Teresa D. Harrison

ECONOMETRIC REVIEWS
2019, VOL. 38, NO. 7, 814–827
https://doi.org/10.1080/07474938.2018.1427486

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

OLS and IV estimation of regression models including endogenous interaction terms

Maurice J. G. Bun[a] and Teresa D. Harrison[b]

[a]Amsterdam School of Economics, Faculty of Economics and Business, University of Amsterdam, Amsterdam, The Netherlands; [b]School of Economics, LeBow College of Business, Drexel University, Philadelphia, PA, USA

ABSTRACT
We analyze a class of linear regression models including interactions of endogenous regressors and exogenous covariates. We show how to generate instrumental variables using the nonlinear functional form of the structural equation when traditional excluded instruments are unknown. We propose to use these instruments with identification robust IV inference. We furthermore show that, whenever functional form identification is not valid, the ordinary least squares (OLS) estimator of the coefficient of the interaction term is consistent and standard OLS inference applies. Using our alternative empirical methods we confirm recent empirical findings on the nonlinear causal relation between financial development and economic growth.

KEYWORDS
Endogeneity; instrumental variables; interaction term; OLS

JEL CLASSIFICATION
C10; C31; C36

1. Introduction

In applied research, it is common to use interaction terms to investigate the multiplicative effect of two variables, labeled x and w, on a dependent variable y. Of primary interest in regression models including interaction terms are the coefficients of those interactions. More specifically, one would like to verify whether the interaction term $x \cdot w$ is significant and economically important and thus should be included in the empirical model.

In this study, we analyze the interaction model in which x is endogenous and w is exogenous. For example, analyzing the returns to schooling one generally regresses wages on education, gender, and other covariates (i.e., ethnicity, age, marital status, etc.). A researcher might interact education and gender in the regression to investigate the gender gap in returns to schooling (Dougherty, 2005). At the same time one may want to correct for endogeneity of education due to selection bias or measurement error. In this case it is expected that the interaction variable, i.e., the product of education and gender, is also an endogenous regressor. Two other examples of empirical studies, in which interactions of endogenous and exogenous regressors appear, are Rajan and Zingales (1998) and Aghion et al. (2005). Both studies analyze the relation between financial development and economic growth, allowing the impact of financial development on growth to be nonlinear.

In the presence of endogenous regressors it is expected that ordinary least squares (OLS) is inconsistent and that instrumental variables (IV) estimation is required instead. Finding valid instruments, however, can prove difficult. For the interaction model, we therefore analyze to what extent one can still produce credible inference without relying on standard exclusion restrictions. First, we show how the functional form of the interaction model naturally leads to alternative instrumental variables, which are functions of the exogenous regressors. To take into account any anomalous effects from weak identification, we propose to use these alternative instruments in combination with a weak instrument

CONTACT Maurice J. G. Bun M.J.G.Bun@uva.nl Amsterdam School of Economics, Faculty of Economics and Business, University of Amsterdam, P.O. Box 15867, 1001 NJ Amsterdam, The Netherlands.

robust test procedure. Second, we establish that, whenever functional form identification fails, the OLS estimator for the coefficient of the interaction term may be consistent and standard inference applies for this parameter of interest.

We demonstrate our theoretical results both through Monte Carlo experiments and by an empirical analysis. The Monte Carlo experiments show the favorable statistical properties of the proposed identification robust IV inference based on the alternative instruments. Also it shows the validity of OLS inference on the interaction term when identification by functional form fails. In addition, we partly reproduce and extend the empirical analysis of Aghion et al. (2005), who analyze the relation between financial development and convergence. In a cross sectional growth regression they test the significance of an interaction effect between initial income and financial development. They allow financial development to be an endogenous regressor. Using our alternative identification procedures, our supplementary empirical results reinforce their conclusion that low financial development makes growth convergence less likely.

In the next section, we describe the interaction model and investigate the asymptotic properties of the IV and OLS estimators and corresponding test procedures. In Section 3, we report Monte Carlo simulation results, while in Section 4 we apply the proposed methods in a growth application. Section 5 concludes.

2. Model and asymptotic properties

2.1. Basic set-up

For ease of exposition, we consider the following model with only one endogenous regressor (labeled x), which interacts with an exogenous regressor (labeled w). Furthermore, there is one additional exogenous regressor (labeled q), which enters in an additive manner only:[1]

$$y_i = \beta_\iota + \beta_w w_i + \beta_x x_i + \beta_{xw} x_i w_i + \beta_q q_i + u_i. \tag{1}$$

One relevant application could be where y is wage, x is schooling, and w is gender. In our application in Section 4, the variables y, x, and w represent country specific growth rates, a measure for the financial development of the country, and log of initial GDP per capita, respectively. The parameter of interest is β_{xw}, i.e., we want to test whether the returns to education is homogeneous or depends on gender or, in our application below, whether the effects of financial development depend on the initial GDP of the country.

To establish the sampling properties of IV and OLS estimators, we make the following assumption regarding the data and errors:

Assumption 1. *The data (y_i, x_i, w_i, q_i) are i.i.d. across i with nonzero finite fourth moments and $E(u_i|w_i, q_i) = 0$.*

Although this simple random sampling assumption rules out most time series applications, it is general enough to allow for conditional heteroskedasticity and non-normality. Note that, we do not specify a particular functional form or relation between the regressors x and w. Hence, x and w can be collinear as is usually the case in applied work.

2.2. IV inference

The structure in (1) and Assumption 1 results in classic endogeneity bias if

$$cov(x_i, u_i) \neq 0. \tag{2}$$

[1] The presence of additional endogenous and exogenous regressors in (1) does not change our theoretical results.

Here, we analyze to what extent we can still identify the coefficients of (1) without having instruments due to standard exclusion restrictions. We supplement the structural equation (1) with the following reduced form for the endogenous regressor x:

Assumption 2a. *The reduced form for x is:*

$$x_i = f\left(w_i, q_i\right) + v_i, \tag{3}$$

with $f(w_i, q_i)$ a nonlinear function of w and q.

Kelejian (1971) demonstrates that using a polynomial approximation to the unknown functional form in (3) provides valid instrumental variables. However, most often this approach has not been used perhaps due to the focus in prior literature on IV estimation of linear models and potential concern about weak instruments. Here, we re-emphasize the usefulness of these instruments, given that due to the particular functional form of the structural equation (1), any nonzero correlation between x and (w, q) will already cause the squares and cross products of the exogenous regressors to be relevant instruments for the interaction term. For this model it is therefore natural to use a second-order polynomial implying the following vector of excluded instruments:[2]

$$z_i = \left[w_i^2 \quad q_i^2 \quad w_i \cdot q_i \quad w_i^2 \cdot q_i \quad w_i \cdot q_i^2\right]'. \tag{4}$$

Our approach does not require any external instruments, but relies exclusively on internal instruments, which are functions of predetermined regressors. In that sense it resembles the use of lagged instruments in time series or panel data applications. Furthermore, it is basically "identification by functional form" where the nonlinear reduced form for the endogenous regressor generates the instruments. It can therefore also be interpreted as a parametric implementation of the semiparametric approach by Escanciano et al. (2016).

Obviously the instruments in (4) are exogenous under Assumption 1 and strong for instrumenting the interaction term.[3] However, their relevance for instrumenting x_i is only guaranteed by a nonlinear reduced form as in Assumption 2a. To take into account possible anomalous effects of weak instruments, we combine the instruments in (4) with recently developed identification robust IV inference. For IV models with i.i.d. errors it is well known that the AR statistic by Anderson and Rubin (1949) testing the joint hypothesis $H_0 : \beta_x = \beta_{x0}, \beta_{xw} = \beta_{xw0}$ is size correct irrespective of the identification strength. Regarding inference on separate coefficients, Guggenberger et al. (2012) show that the subset AR tests for $H_0 : \beta_x = \beta_{x0}$ and $H_0 : \beta_{xw} = \beta_{xw0}$ have correct asymptotic size too. Furthermore, Kleibergen and Mavroeidis (2009) show that similar results hold for the GMM extension of the AR statistic by Stock and Wright (2000), which is robust to heteroskedasticity.

One problem with the AR statistic is that the corresponding AR confidence intervals for β_x and β_{xw} may be inaccurate. Davidson and MacKinnon (2014a) analyze AR confidence intervals in detail and conclude that often they are misleading. In the empirical analysis, we therefore also consider an improvement of the standard Wald-based confidence intervals by exploiting the restricted efficient (RE) bootstrap. Davidson and MacKinnon (2014b) show that RE bootstrap confidence intervals based on IV Wald t-statistics are quite reliable.

Summarizing, our approach for the interaction model (1) is the combination of identification by functional form and weak instrument robust inference. The former provides the alternative set of instruments, while the latter preserves asymptotically size correct inference in all cases.

[2] When w_i is a dummy variable the powers of w_i do not provide additional identifying information. In that special case one can exploit $z_i = \left[q_i^2 \quad w_i \cdot q_i \quad w_i \cdot q_i^2\right]'$, see Section 3 for this example illustrated in our Monte Carlo study.

[3] To economize on the number of instruments one might want to use only a subset of the instruments e.g. omitting the cross terms.

2.3. OLS inference

When the reduced form is linear, Assumption 2a does not hold and the "identification by functional form" argument does not apply. However, we show below that in this case, OLS will be consistent for the coefficient of the interaction term. We do so under the following alternative assumption:

Assumption 2b. *The reduced form is linear:*

$$x_i = \pi_\iota + \pi_w w_i + \pi_q q_i + v_i, \tag{5}$$

with

$$E(u_i v_i | w_i, q_i) = E(u_i v_i) = \sigma_{uv}. \tag{6}$$

Under Assumption 2b the reduced form coefficients for the instruments in (4) are zero. Therefore, the concentration parameter is zero, even asymptotically, leading to irrelevant instruments for x. Although weak identification robust inference is still valid, it will lead to large and possibly unbounded confidence intervals (Dufour, 1997).

Assumption 2b is equivalent to the assumptions underlying Theorem 1 of Lewbel (2012). It implies that $w_i v_i$ and $q_i v_i$ are valid alternative instruments as they are uncorrelated with the structural error u_i. Lewbel (2012) notes that the strength of these additional instruments depends crucially on the covariances between $w_i v_i$ and $q_i v_i$ with v_i. Therefore, identification is achieved whenever reduced form errors exhibit heteroskedasticity such that $Cov(w_i, v_i^2)$ and $Cov(q_i, v_i^2)$ are nonzero. We don't discuss the resulting inference here, however, as details are already provided in Lewbel (2012). Instead, we abstract away from identification by heteroskedasticity and therefore assume the following:

Assumption 2c.

$$E\left(v_i^2 | w_i, q_i\right) = E\left(v_i^2\right) = \sigma_v^2.$$

We analyze the properties of the OLS estimator under Assumption 2b and 2c, i.e., when identification by both functional form and heteroskedasticity do not apply for the interaction model. We abstract without loss of generalization from the inclusion of additional exogenous regressors q_i.[4] Stacking the observations ($i = 1, \ldots, n$), we then get

$$y = X\beta + u, \tag{7}$$

where $y = (y_1, \ldots, y_n)'$, $X = (X_1', \ldots, X_n')'$ with $X_i = \begin{bmatrix} 1 & w_i & x_i & x_i w_i \end{bmatrix}'$, $\beta = (\beta_\iota, \beta_w, \beta_x, \beta_{xw})'$ and $u = (u_1, \ldots, u_n)'$. The OLS estimator of the full parameter vector β is equal to:

$$\hat{\beta} = (X'X)^{-1} X'y. \tag{8}$$

Taking the probability limit, we have

$$\text{plim}\hat{\beta} = \beta + \Sigma_{XX}^{-1} \Sigma_{Xu}, \tag{9}$$

where we defined $\Sigma_{XX} = \text{plim}\frac{1}{n} X'X$ and $\Sigma_{Xu} = \text{plim}\frac{1}{n} X'u$. The vector $\Sigma_{XX}^{-1} \Sigma_{Xu}$ is the OLS inconsistency.

To derive the limiting distribution of the OLS estimator, rewrite model (1) as:

$$y_i = X_i' \beta_* + \varepsilon_i, \tag{10}$$

[4] The presence of additional exogenous regressors does not change the theoretical results as they enter in a linear fashion in both structural equation (1) and, under Assumption 2b, also in the reduced form equation for x. The analysis below therefore holds exactly when we replace y, w, and x by the residuals of their projection on these additional exogenous regressors.

with $\beta_* = \beta + \Sigma_{XX}^{-1}\Sigma_{Xu}$ the pseudo-true value and

$$\varepsilon_i = u_i - \Sigma_{Xu}'\Sigma_{XX}^{-1}X_i, \tag{11}$$

such that $E[X_i\varepsilon_i] = 0$. The OLS estimator (8) is simply a method of moments estimator exploiting the following moment equation:

$$E\left[X_i\left(y_i - X_i'\beta\right)\right] = 0. \tag{12}$$

These moment conditions are satisfied when $\beta = \beta_*$. Standard asymptotic theory for method of moments estimators then gives the following large sample distribution of the OLS estimator:

Lemma 1. *Given model (1) and Assumption 1, the large sample distribution of the OLS estimator (8) is:*

$$\sqrt{n}\left(\hat{\beta} - \beta_*\right) \xrightarrow{d} \mathcal{N}(0, V), \tag{13}$$

where

$$V = \Sigma_{XX}^{-1}\Sigma_{X\varepsilon\varepsilon X}\Sigma_{XX}^{-1}, \tag{14}$$

with

$$\Sigma_{X\varepsilon\varepsilon X} = \text{plim}\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 X_i X_i'. \tag{15}$$

From Lemma 1 it can be seen that, although normally distributed, the limiting distribution of the OLS estimator is centered around its pseudo-true value $\beta_* = \beta + \Sigma_{XX}^{-1}\Sigma_{Xu}$, but with a standard sandwich-type expression for the asymptotic variance. Therefore, valid OLS inference results whenever the OLS inconsistency $\Sigma_{XX}^{-1}\Sigma_{Xu}$ is zero. Now, we can show:

Proposition 1. *Under Assumptions 1, 2b, and 2c, the inconsistency of the OLS estimator of model (1) equals:*

$$\Sigma_{XX}^{-1}\Sigma_{Xu} = \frac{\sigma_{uv}}{\sigma_v^2}\begin{bmatrix} -\pi_\iota \\ -\pi_w \\ 1 \\ 0 \end{bmatrix}. \tag{16}$$

Proof. see the Appendix.

Lemma 1 and Proposition 1 together imply the following limiting distribution for the OLS estimator of the interaction coefficient β_{xw}:

$$\sqrt{n}\left(\hat{\beta}_{xw} - \beta_{xw}\right) \xrightarrow{d} \mathcal{N}\left(0, V_{xw}\right), \tag{17}$$

where V_{xw} is the diagonal element of V corresponding to the interaction term. In other words, we have that, even if we have an endogenous regressor x, the OLS estimator of the coefficient β_{xw} is consistent and standard heteroskedasticity-robust OLS inference applies.

2.4. Discussion

Regarding the parameter of interest, i.e., the coefficient of the interaction term β_{xw}, the proposed IV procedure and the OLS estimator are to some extent complementary to each other. In other words, whenever identification by functional form is weak, and therefore Assumption 2a is not valid, the linear functional form of Assumption 2b holds. Under the additional restrictions in Assumptions 2b and 2c, OLS inference is then valid. One can determine the validity of Assumption 2a and 2b simply by performing Wald tests for whether polynomials and cross products of w and q are significant in the reduced form equation. In addition, standard heteroskedasticity tests can be used to verify Assumption 2c.

Under Assumptions 2b and 2c the coefficient of the interaction term is identified by OLS, irrespective of the further nature of the distribution of x and w. In particular, w could be discrete or continuous. An obvious special case is that of joint normality of x_i, w_i, and u_i. In this case, the conditional distribution of (x_i, u_i) is normal with conditional covariance not depending on w_i. Another example is independence of x_i and w_i as in the treatment regression model analyzed in Nizalova and Murtazashvili (2016). Lewbel (2012) discusses other examples where Assumption 2b is satisfied, i.e., classical measurement error and unobserved single factor models. Unlike a linear structural equation, however, Assumption 2b will fail when the endogeneity in the interaction model (1) is due to simultaneous causality. The reason is that the nonlinear simultaneity in the structural equation carries over to a nonlinear reduced form. But in this case functional form identification and the resulting IV estimator will apply again.

It should be noted that the OLS consistency is restricted to β_{xw} only and not the full marginal effect of x on y (i.e., $\beta_x + \beta_{xw}w$) because the OLS estimator of β_x is inconsistent always. However, OLS inference can serve as a pretesting diagnostic for whether the interaction term should be in the structural equation. Furthermore, the marginal effect of w on y can be consistently estimated whenever x and w are independent as is the case in the treatment regression model discussed above. Finally, in many studies the coefficient of interest is precisely the interaction coefficient as in our example in Section 4 (i.e., Aghion et al., 2005). Well known results for models including interaction terms (Allison, 1977; Braumoeller, 2004) show that the lower-order coefficients β_w and β_x are not of direct interest as their values can be manipulated easily by scaling of the data. In contrast, the coefficient of the interaction term is invariant to scaling, which extends to testing hypotheses about β_{xw} (Allison, 1977).

3. Monte Carlo simulations

We performed a number of Monte Carlo experiments to verify the accuracy of the proposed IV and OLS inference methods for the interaction model. The IV procedure based on functional form identification uses (4) as instrument set and we refer to it as IVF.

We simulate the finite sample distributions of the IVF and OLS coefficient estimators. Apart from analyzing coefficient bias, we also report actual rejection probabilities of corresponding Wald t-statistics. Moreover, regarding the IVF procedure we additionally report actual rejection frequencies of the Anderson–Rubin (AR) test. We report both the test of the joint null hypothesis $H_0 : \beta_x = \beta_{x0}$, $\beta_{xw} = \beta_{xw0}$ as well as the subset AR test procedures proposed by Guggenberger et al. (2012) for testing $H_0 : \beta_x = \beta_{x0}$ and $H_0 : \beta_{xw} = \beta_{xw0}$ separately.

Data for y and x were generated by (1) and (3), respectively, assuming various distributions for the exogenous regressors and errors. For the structural equation, we choose $\beta_\iota = \beta_w = \beta_x = \beta_{xw} = \beta_q = 1$. Regarding the reduced form (3) for x we experimented with linear and quadratic functional forms, i.e., we generate x according to:

$$x_i = \pi_\iota + \pi_1 w_i + \pi_2 w_i^2 + \pi_3 q_i + \pi_4 q_i^2 + v_i. \tag{18}$$

The quadratic and linear reduced form experiments correspond to $\pi_2 = \pi_4 = 0.5$ and $\pi_2 = \pi_4 = 0$, respectively, while in both experiments we choose $\pi_\iota = \pi_1 = \pi_3 = 1$.[5] We generate normally[6]

[5]We experimented with alternative values for these reduced form parameters resulting in qualitatively similar simulation results.
[6]Unreported simulation experiments with errors generated by other distributions than normal show similar results.

distributed errors $(u_i, v_i) \sim$i.i.n.$(0, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 1 & \rho_{uv} \\ \rho_{uv} & 1 \end{pmatrix}.$$

The parameter ρ_{uv} determines the degree of endogeneity. Furthermore, we choose $w_i \sim$i.i.n.$(0, 1)$, $q_i \sim$i.i.n.$(0, 1)$, and $n = 100$ in the experiments. All simulation results are based on 20,000 replications.

In the experiment with the quadratic reduced form ($\pi_2 \neq 0$) Assumption 2b is not satisfied, hence our theoretical results predict that OLS inference fails. However, AR inference based on the instruments in (4) is size correct with nontrivial power. Alternatively, under Assumption 2b ($\pi_2 = 0$) identification by functional form fails and IVF AR statistics have flat power curves. However, Proposition 1 shows that OLS in this case is consistent for the interaction term, and its t-statistic corresponding to β_{xw} is asymptotically size correct.

In Tables 1 and 2, we report simulation results for the quadratic and linear reduced form. To save space we only show results for the coefficients of the endogenous regressors. From Table 1, we can see that the OLS estimators for both β_x and β_{xw} are indeed biased in the quadratic reduced form design. The preferred method is IVF and biases in the IVF coefficient estimators are substantially smaller, as expected. Both Wald and AR based IVF inference is accurate, i.e., actual rejection frequencies are close to the nominal level of 5%. Furthermore, Table 2 shows that AR based inference also has reasonably large power.

In the linear reduced form design results are quite different. In this case the OLS coefficient estimator of β_{xw} is consistent resulting in negligible bias. Estimation of β_x by OLS still results in large bias, but now also the IVF estimator exhibits large bias for this coefficient. The reason is that under Assumption 2b, i.e., a linear reduced form, the instruments are irrelevant for x. OLS and IVF coefficient biases (and lack of bias as it may be) carry over to Wald t-tests: the actual rejection frequency for testing β_{xw} is close to the nominal level, while that for testing β_x can be way off. Inference based on the subset AR statistics is

Table 1. Coefficient bias and test size.

	Quadratic reduced form				Linear reduced form			
	OLS		IVF		OLS		IVF	
	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}
Bias	0.285	−0.049	0.016	−0.004	0.500	−0.000	0.501	0.000
sd	0.080	0.050	0.131	0.073	0.090	0.048	0.616	0.068
rp Wald *t*	91.65	18.32	5.82	4.96	99.94	5.55	14.18	2.65
rp AR joint			7.39				7.39	
rp AR subset			4.38	4.90			4.46	0.83

Note: Based on 20,000 MC replications; rp is actual rejection % of nominal 5% tests.

Table 2. Rejection frequencies under null and alternative hypotheses

	Quadratic reduced form				Linear Reduced form			
	OLS		IVF		OLS		IVF	
Δ	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}
−0.5	100.00	100.00	46.55	94.55	100.00	100.00	4.51	4.61
−0.4	99.99	99.96	35.05	89.31	100.00	100.00	4.49	4.38
−0.3	99.99	98.69	23.25	75.70	100.00	99.92	4.51	4.15
−0.2	99.88	80.11	12.67	46.86	100.00	95.78	4.43	3.47
−0.1	98.88	18.14	6.37	14.83	100.00	52.99	4.44	1.78
0	91.65	18.32	4.38	4.90	99.94	5.55	4.46	0.83
0.1	60.56	78.29	7.21	12.48	98.90	52.96	4.37	1.89
0.2	18.85	98.14	17.73	32.10	89.71	95.97	4.45	3.44
0.3	6.00	99.84	38.33	51.18	58.39	99.87	4.59	4.17
0.4	29.30	99.99	61.81	65.66	19.50	99.98	4.57	4.51
0.5	73.08	100.00	78.89	75.19	5.05	100.00	4.52	4.57

Note: Based on 20,000 MC replications; Δ indicates difference between true value and hypothesized value; Regarding IVF only the AR subset statistics are shown.

actually somewhat conservative under H_0, which is in line with the results of Guggenberger et al. (2012). Furthermore, Table 2 shows that power curves are flat indicating the lack of identification. OLS inference on β_{xw}, however, is excellent in this case. There is virtually no size distortion, while power is high already for nearby alternative hypotheses.

In Table 3, we report simulation results for a larger sample size $n = 250$. Compared with Table 1 the magnitude of OLS and IVF coefficient biases stay the same. OLS standard deviations decrease as a result of the larger sample size, but not for IVF when the reduced form is linear. In this case the instruments are irrelevant for x, hence the IVF estimator is inconsistent. Due to the increased precision of the OLS estimator, size distortions are larger whenever the OLS estimator is inconsistent.

In Table 4, we go back to the original sample size $n = 100$, but specify w_i to be a binary variable with 50% zeros and 50% ones. Because in this design $w_i = w_i^2$, the reduced form (18) is always linear in w_i (hence it can only be nonlinear in q_i). Therefore, the OLS coefficient estimator of β_{xw} is consistent. Table 4 indeed reports negligible bias of the OLS coefficient estimator of β_{xw} as well as the corresponding Wald t-test. Regarding IVF, when w_i is a dummy variable the powers of w_i do not provide additional identifying information. In this case, we use $z_i = \begin{bmatrix} q_i^2 & w_i \cdot q_i & w_i \cdot q_i^2 \end{bmatrix}'$ instead of (4), see also footnote 2. The IVF simulation results in Table 4 show that biases are indeed very small when the reduced form is quadratic. In the linear case, however, the instruments are again irrelevant for x and identification by functional form breaks down. Also note that the AR tests are somewhat conservative in this design, although asymptotically they have the correct size (Guggenberger et al., 2012). This discrepancy is due to the relatively small sample size of $n = 100$ in this simulation experiment.

4. Economic growth and financial development

Aghion et al. (2005) develop a theory implying that economic growth convergence depends on the level of financial development. They test their theory in a cross country growth regression including an interaction term between initial GDP per capita and an indicator of financial development. Sample size is $n = 71$ countries. In our notation y_i is the average growth rate of GDP per capita in the period 1960–1995, w_i is initial (1960) per capita GDP and x_i is the average level of financial development. Their

Table 3. Coefficient bias and test size for $n = 250$.

	Quadratic reduced form				Linear reduced form			
	OLS		IVF		OLS		IVF	
Δ	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}
Bias	0.278	−0.047	0.006	−0.001	0.500	−0.000	0.497	0.000
sd	0.050	0.030	0.078	0.042	0.055	0.028	0.612	0.040
rp Wald t	99.61	36.95	5.33	4.63	100.00	5.16	13.72	2.51
rp AR joint			5.98				5.98	
rp AR subset			4.81	5.08			4.79	0.57

Note: See Table 1.

Table 4. Coefficient bias and test size in case of binary w.

	Quadratic reduced form				Linear reduced form			
	OLS		IVF		OLS		IVF	
Δ	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}
Bias	0.345	−0.000	−0.003	0.001	0.500	0.001	0.499	0.005
sd	0.108	0.133	0.195	0.188	0.110	0.127	1.813	0.376
rp Wald t	85.90	5.42	4.57	4.22	98.81	4.85	4.74	1.55
rp AR joint			1.72				1.72	
rp AR subset			0.60	0.68			0.63	0.03

Note: see Table 1. Sample size n=100.

specifications include different sets of control variables q_i (labeled "empty", "policy," and "full").[7] The data are taken from Levine et al. (2000) and include four different measures of financial development ("private credit," "liquid liabilities," "bank assets," and "commercial-central bank").

Aghion et al. (2005) conjecture that financial development is an endogenous regressor because of feedback from growth to finance, or because of relevant omitted variables. They acknowledge that the interaction between financial development and initial income may be an endogenous regressor too. They follow La Porta et al. (1997, 1998) and use legal origin as a source of exogenous variation in financial development. Legal origin is quantified by three binary indicators measuring French, English, and German traditions.

Table 5 reports IV and OLS estimation results using private credit as the measure of financial development.[8] The first specification does not include any further control variables; Aghion et al. (2005) consider this specification to be representative of their main result. The second and third set of results use their policy and full set of control variables, respectively.

Aghion et al. (2005) use the IV estimator exploiting external legal origin instruments (labeled AHM), and the reported coefficient estimates are in columns (1), (4), and (7) of Table 5. Note that Aghion et al. (2005) used homoskedasticity-only standard errors, which we use as well for the IV estimation.[9]

The IV estimation results exploiting functional form identification (labeled IVF) are in columns (5) and (8). Regarding the specification without control variables it is not defined as we have only one instrument (w^2) for two endogenous regressors.[10] For the regressions with policy and full sets of control variables we used (4), but to economize on the number of instruments we omitted the cross products and only used the squared values of all exogenous variables as instruments. According to the Sargan test these alternative instruments are exogenous for the policy specification. For the full set of control variables we find rejection by the Sargan test of the validity of the additional squared instruments. Further inspection of the results reveals that squared ethnic diversity actually should be included as a further control variable instead of an instrument. Doing so produces the empirical results in column (8). Finally, OLS results are reported in columns (3), (6), and (9).

Table 5 shows that, although coefficient estimates sometimes differ in magnitude, the pattern of the AHM, IVF, and OLS results is similar across specifications. Regarding the interaction coefficient, the

Table 5. Empirical results for GDP growth and private credit.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Empty			Policy			Full		
	AHM	IVF	OLS	AHM	IVF	OLS	AHM	IVF	OLS
β_x	−0.015		−0.012	−0.013	−0.008	−0.009	−0.016	−0.004	−0.006
	(0.015)		(0.008)	(0.018)	(0.016)	(0.007)	(0.018)	(0.015)	(0.008)
β_{xw}	−0.061		−0.048	−0.063	−0.052	−0.042	−0.063	−0.044	−0.037
	(0.011)		(0.008)	(0.011)	(0.011)	(0.009)	(0.012)	(0.011)	(0.011)
CD	6.28			3.87	4.16		3.64	3.07	
KP	26.57			19.99	20.99		19.98	22.57	
	(0.00)			(0.00)	(0.00)		(0.00)	(0.00)	
Sargan	3.13			2.06	1.51		3.00	1.82	
	(0.54)			(0.73)	(0.83)		(0.56)	(0.94)	

Note: Numbers in parentheses are standard errors (below estimates) or *p*-values (below test statistics); IV and OLS standard errors are homoskedasticity-only and heteroskedasticity-robust, respectively; Empty, Policy, and Full refer to the set of control variables; Financial development measure (x) is private credit; CD and KP are Cragg–Donald and Kleibergen–Paap statistics, respectively. Regarding CD critical values are 9.48/6.08/4.78 for 10/20/30% relative IV bias, see Stock and Yogo (2005).

[7]The policy control variables are average years of schooling, government size, inflation, black market premium, and trade openness. The full conditioning set is the policy set plus indicators for revolution and coups, political assassinations, and ethnic diversity.

[8]Private credit is the preferred measure of financial development according to Aghion et al. (2005).

[9]Tests for heteroskedasticity do not reject the null hypothesis of homoskedasticity. However, as warranted by Lemma 1, we report heteroskedasticity-robust OLS standard errors.

[10]One can simply add w^3 to meet the order condition. For brevity we refrained here from using third or higher order terms, also because this would lead to a proliferation of instruments in the models with control variables.

Table 6. 95% Confidence intervals.

	Wald		RE bootstrap Wald		AR	
	β_x	β_{xw}	β_x	β_{xw}	β_x	β_{xw}
(1)	(−0.045, 0.015)	(−0.083, −0.039)	(−0.085, 0.015)	(−0.095, −0.045)	(−0.071, 0.038)	(−0.108, −0.032)
(2)						
(3)	(−0.027, 0.003)	(−0.064, −0.033)				
(4)	(−0.049, 0.022)	(−0.085, −0.040)	$(-\infty, \infty)$	(−0.095, −0.035)	(−0.098, 0.077)	(−0.120, −0.026)
(5)	(−0.040, 0.024)	(−0.073, −0.030)	(−0.052, 0.025)	(−0.083, −0.020)	(−0.075, 0.071)	(−0.103, −0.014)
(6)	(−0.023, 0.005)	(−0.061, −0.023)				
(7)	(−0.052, 0.020)	(−0.087, −0.039)	$(-\infty, \infty)$	(−0.096, −0.027)	(−0.114, 0.071)	(−0.132, −0.026)
(8)	(−0.032, 0.024)	(−0.064, −0.023)	(−0.045, 0.049)	(−0.092, −0.013)	(−0.095, 0.092)	(−0.111, 0.011)
(9)	(−0.022, 0.009)	(−0.059, −0.016)				

Note: Row labeling refers to estimated specifications in Table 5. Number of bootstrap replications is 10,000.

AHM coefficient estimate is largest (in absolute value) followed by IVF and OLS, while the opposite ranking holds for the standard errors. Conventional inference based on Wald t-statistics indicate that the interaction coefficient is significantly different from zero.[11]

Although the Kleibergen and Paap (2006) rank statistic does reject the null hypothesis of underidentification in all specifications of Table 5, the reported Cragg and Donald (1993) statistics indicate the possibility of weak instruments as demonstrated by the test procedure developed by Stock and Yogo (2005). They develop a weak instrument test based on the Cragg–Donald (CD) statistic reported in Table 5, whose value can be related to bias of the IV estimator relative to the bias of the OLS estimator. Applying the critical values of Stock and Stock and Yogo (2005) to the reported CD statistic[12] we find that the null hypothesis of 30% relative bias cannot be rejected for all IV specifications except column (1), indicating a weak instrument problem for most of the specifications.

The subset RE bootstrap Wald and AR confidence intervals take this into account, and indeed they are larger than their Wald counterparts as can be seen in Table 6. When bounded, RE bootstrap Wald based confidence intervals are shorter than AR based confidence intervals. The length of the AR intervals depends on the value of the Sargan statistic corroborating the results of Davidson and MacKinnon (2014a). Their analysis reveals that the AR confidence intervals tend to be too long when the Sargan statistic is small and vice versa. In this application the dependence on the value of the Sargan statistic makes the AR confidence intervals therefore likely to be misleadingly long. The RE bootstrap Wald based confidence intervals can be a viable alternative. Irrespective of the method, however, the coefficient of the interaction term is significantly different from zero in almost all cases, confirming the main result of Aghion et al. (2005) of a nonlinear relation between financial development and growth.

5. Concluding remarks

In this study, we have analyzed IV and OLS inference for regression models including interactions between endogenous regressors and exogenous covariates. First, we show how to achieve identification without relying on standard exclusion restrictions, but merely by exploiting the nonlinear functional form. We combine the alternative instrument set with identification robust IV statistics to obtain size correct IV inference. Second, we show that endogeneity bias can be reduced to zero for the OLS estimator as far as the interaction term is concerned. Whenever IV based inference procedures fail, we show that the OLS estimator of the coefficient of the interaction term is consistent, and that standard OLS inference applies.

Monte Carlo experiments corroborate our theoretical findings. In particular, we show that identification robust IV inference based on the alternative instrument set is size correct with nontrivial power

[11] We checked Assumptions 2a and 2b by performing Wald tests for whether nonlinearities are present in the reduced form. For the empty specification we don't find evidence of a nonlinear reduced form, hence OLS inference on β_{xw} is valid. For the policy and full specifications, however, we find significant nonlinearities and IVF inference is preferred.

[12] Critical values are 9.48/6.08/4.78 for 10/20/30% relative IV bias, see Stock and Yogo (2005).

in a variety of cases. In addition, we show that in finite samples OLS inference on the interaction term is accurate when identification by functional form fails.

We partly reproduce and extend the empirical analysis of Aghion et al. (2005), who analyze the interaction between financial development and growth convergence. We provide further support for their instrument set choice, but at the same time show that identification by functional form or OLS can be equally valid. Our supplementary empirical results reinforce their conclusion that low financial development makes growth convergence less likely.

Our results are derived under quite general conditions, allowing for continuous and discrete interaction terms, correlation between endogenous and exogenous regressors, conditional heteroskedasticity, and non-normality. Thus, in interaction models the researcher can always perform valid statistical inference for the interaction term without the use of standard IV exclusion restrictions. Our proposed methods are particularly useful in applications where no external instrumental variables are available or their validity is questionable. They can also be used in addition to classical instruments to improve efficiency.

Appendix

Proof of Proposition 1. The components of the OLS estimation error are:

$$X_i u_i = \begin{pmatrix} u_i \\ w_i u_i \\ x_i u_i \\ x_i w_i u_i \end{pmatrix}, \; X_i X_i' = \begin{pmatrix} 1 & w_i & x_i & x_i w_i \\ w_i & w_i^2 & x_i w_i & x_i w_i^2 \\ x_i & x_i w_i & x_i^2 & x_i^2 w_i \\ x_i w_i & x_i w_i^2 & x_i^2 w_i & x_i^2 w_i^2 \end{pmatrix}. \tag{19}$$

Assumption 1 implies

$$\Sigma_{XX} \equiv \text{plim} \frac{1}{n} X'X = \lim \frac{1}{n} \sum_{i=1}^{n} E\left[X_i X_i'\right] = E\left[X_i X_i'\right], \tag{20}$$

and

$$\Sigma_{Xu} \equiv \text{plim} \frac{1}{n} X'u = \lim \frac{1}{n} \sum_{i=1}^{n} E\left[X_i u_i\right] = E\left[X_i u_i\right], \tag{21}$$

noting that the last equality occurs under the i.i.d. assumption as the values of $E\left[X_i X_i'\right]$ and $E\left[X_i u_i\right]$ do not depend on i. Hence, regarding the OLS inconsistency $\Sigma_{XX}^{-1}\Sigma_{Xu}$ we can evaluate the separate elements of $E\left[X_i X_i'\right]$ and $E\left[X_i u_i\right]$:

$$E\left[X_i X_i'\right] = \begin{pmatrix} 1 & E[w_i] & E[x_i] & E[x_i w_i] \\ E[w_i] & E\left[w_i^2\right] & E[x_i w_i] & E\left[x_i w_i^2\right] \\ E[x_i] & E[x_i w_i] & E\left[x_i^2\right] & E\left[x_i^2 w_i\right] \\ E[x_i w_i] & E\left[x_i w_i^2\right] & E\left[x_i^2 w_i\right] & E\left[x_i^2 w_i^2\right] \end{pmatrix}, \; E\left[X_i u_i\right] = \begin{pmatrix} 0 \\ 0 \\ E[x_i u_i] \\ E[x_i w_i u_i] \end{pmatrix}. \tag{22}$$

Assumptions 2b and 2c furthermore imply for the nonzero elements in (22)

$$\begin{aligned} E[x_i] &= \pi_\iota + \pi_w E[w_i] \\ E[x_i w_i] &= \pi_\iota E[w_i] + \pi_w E\left[w_i^2\right] \\ E\left[x_i w_i^2\right] &= \pi_\iota E\left[w_i^2\right] + \pi_w E\left[w_i^3\right] \\ E\left[x_i^2\right] &= \pi_\iota^2 + \pi_w^2 E\left[w_i^2\right] + \sigma_v^2 + 2\pi_\iota \pi_w E[w_i] \end{aligned}$$

$$E\left[x_i^2 w_i\right] = \pi_\iota^2 E\left[w_i\right] + \pi_w^2 E\left[w_i^3\right] + \sigma_v^2 E\left[w_i\right] + 2\pi_\iota \pi_w E\left[w_i^2\right]$$

$$E\left[x_i^2 w_i^2\right] = \pi_\iota^2 E\left[w_i^2\right] + \pi_w^2 E\left[w_i^4\right] + \sigma_v^2 E\left[w_i^2\right] + 2\pi_\iota \pi_w E\left[w_i^3\right]$$

$$E\left[x_i u_i\right] = \sigma_{uv}$$

$$E\left[x_i w_i u_i\right] = \sigma_{uv} E\left[w_i\right]$$

We have that

$$\Sigma_{XX}^{-1} = \frac{1}{\det\left(\Sigma_{XX}\right)} adj(\Sigma_{XX}), \tag{23}$$

where the transpose of $adj(\Sigma_{XX})$ is the matrix of cofactors of Σ_{XX}. Because Σ_{Xu} has only the third and fourth elements nonzero, for the evaluation of $\Sigma_{XX}^{-1}\Sigma_{Xu}$ we only need cofactors corresponding to the third and fourth columns of Σ_{XX}. Denoting with c_{ij} the cofactor of entry d_{ij} in matrix Σ_{XX}, we have:

$$\begin{aligned}
c_{13} = \det &\begin{pmatrix} E\left[w_i\right] & E\left[w_i^2\right] & E\left[x_i w_i^2\right] \\ E\left[x_i\right] & E\left[x_i w_i\right] & E\left[x_i^2 w_i\right] \\ E\left[x_i w_i\right] & E\left[x_i w_i^2\right] & E\left[x_i^2 w_i^2\right] \end{pmatrix} \\
= &\ \pi_w \sigma_v^2 E(w_i)\left(E(w_i^2)^2 - E(w_i)E(w_i^3)\right) + \pi_\iota \sigma_v^2 \left(E(w_i)^2 - E(w_i^2)\right) E(w_i^2) \\
&+ \pi_\iota \pi_w^2 \left(E(w_i^2)^3 - 2E(w_i)E(w_i^2)E(w_i^3) + E(w_i^3)^2\right) \\
&+ \pi_\iota^3 \left(E(w_i)^2 - E(w_i^2)\right) E(w_i^2)(\pi_w^2 E(w_i^4) - 1).
\end{aligned} \tag{24}$$

In a similar way, we get:

$$c_{14} = \pi_\iota \sigma_v^2 E(w_i)\left(E(w_i^2) - E(w_i)^2\right) + \pi_w \sigma_v^2 \left(E(w_i)E(w_i^3) - E(w_i^2)^2\right), \tag{25}$$

$$\begin{aligned}
c_{23} = &\ \pi_w \sigma_v^2 \left(E(w_i)E(w_i^3) - E(w_i^2)^2\right) + \pi_\iota \sigma_v^2 \left(E(w_i^2) - E(w_i)^2\right) E(w_i) \\
&+ \pi_w^3 \left(E(w_i^2)^3 - 2E(w_i)E(w_i^2)E(w_i^3) + E(w_i^3)^2\right) \\
&+ \pi_\iota^2 \pi_w \left(E(w_i)^2 - E(w_i^2)\right) E(w_i^2)(\pi_w^2 E(w_i^4) - 1),
\end{aligned} \tag{26}$$

$$c_{24} = -\pi_\iota \sigma_v^2 \left(E(w_i^2) - E(w_i)^2\right) - \pi_w \sigma_v^2 \left(E(w_i)^3 - 2E(w_i)E(w_i^2) + E(w_i^3)\right), \tag{27}$$

$$\begin{aligned}
c_{33} = &\ \sigma_v^2 E(w_i^2)\left(E(w_i^2) - E(w_i)^2\right) \\
&- \pi_w^2 \left(E(w_i^2)^3 - 2E(w_i)E(w_i^2)E(w_i^3) + E(w_i^3)^2\right) \\
&- \pi_\iota^2 \left(E(w_i)^2 - E(w_i^2)\right) E(w_i^2)(\pi_w^2 E(w_i^4) - 1),
\end{aligned} \tag{28}$$

$$c_{34} = -\sigma_v^2 E(w_i)\left(E(w_i^2) - E(w_i)^2\right), \tag{29}$$

$$c_{43} = -\sigma_v^2 E(w_i)\left(E(w_i^2) - E(w_i)^2\right), \tag{30}$$

$$c_{44} = \sigma_v^2 \left(E(w_i^2) - E(w_i)^2\right), \tag{31}$$

and also

$$\begin{aligned}\det(\Sigma_{XX}) = {} & \sigma_v^4 \left(E(w_i^2) - E(w_i)^2\right)^2 \\ & - \sigma_v^2 \pi_w^2 \left(E(w_i^2)^3 - 2E(w_i)E(w_i^2)E(w_i^3) + E(w_i^3)^2\right) \\ & - \sigma_v^2 \pi_\iota^2 \left(E(w_i)^2 - E(w_i^2)\right) E(w_i^2)(\pi_w^2 E(w_i^4) - 1). \end{aligned} \tag{32}$$

Substituting the expressions (24)–(32) into (22) yields

$$\begin{aligned}\Sigma_{XX}^{-1}\Sigma_{Xu} &= \frac{\sigma_{uv}}{\det(\Sigma_{XX})} \begin{bmatrix} c_{13} + E(w_i)c_{14} \\ c_{23} + E(w_i)c_{24} \\ c_{33} + E(w_i)c_{34} \\ c_{43} + E(w_i)c_{44} \end{bmatrix} \\ &= \frac{\sigma_{uv}}{\sigma_v^2} \begin{bmatrix} -\pi_\iota \\ -\pi_w \\ 1 \\ 0 \end{bmatrix}, \end{aligned} \tag{33}$$

which completes the proof.

Acknowledgment

This article greatly benefited from comments made by the Editor and two anonymous referees. We also like to thank Otilia Boldea, Stephane Bonhomme, Peter Boswijk, Simon Broda, Artūras Juodis, Jan Kiviet, Frank Kleibergen, James MacKinnon, Andrew Pua, and seminar participants for helpful comments.

Funding

The research of the author has been funded by the NWO Vernieuwingsimpuls research grant nr. 016.115.320.

References

Aghion, P., Howitt, P., Mayer-Foulkes, D. (2005). The effect of financial development on convergence: Theory and evidence. *The Quarterly Journal of Economics* 120(1):173–222.

Allison, P. D. (1977). Testing for interaction in multiple regression. *American Journal of Sociology* 83:144–153.

Anderson, T. W., Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics* 20:46–63.

Braumoeller, B. F. (2004). Hypothesis testing and multiplicative interaction terms. *International Organization* 58:807–820.

Cragg, J. G., Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory* 9:222–240.

Davidson, R., MacKinnon, J. G. (2014a). Confidence sets based on inverting Anderson-Rubin tests. *Econometrics Journal* 17:S39–S58.

Davidson, R., MacKinnon, J. G. (2014b). Bootstrap confidence sets with weak instruments. *Econometric Reviews* 33: 651–675.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men? *Journal of Human Resources* 40:969–988.

Dufour, J.-M. (1997). Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* 65:1365–1387.

Escanciano, J.C., Jacho-Chávez, D., Lewbel, A. (2016). Identification and estimation of semiparametric two step models. *Quantitative Economics* 7:561–589.

Guggenberger, P., Kleibergen, F., Mavroeidis, S., Chen, L. (2012). On the asymptotic sizes of subset Anderson-Rubin and Lagrange Multiplier tests in linear instrumental variables regression. *Econometrica* 80:2649–2666.

Kelejian, H. H. (1971). Two-Stage Least Squares and econometric systems linear in parameters but nonlinear in the endogenous variables. *Journal of the American Statistical Association* 66:373–374.

Kleibergen, F., Mavroeidis, S. (2009). Weak instrument robust tests in GMM and the new Keynesian Phillips curve. *Journal of Business and Economic Statistics* 27:293–311.

Kleibergen, F., Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133:97–126.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. W. (1997). Legal determinants of external finance. *Journal of Finance* 52:1131–1150.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. W. (1998). Law and finance. *Journal of Political Economy* 106: 1113–1155.

Levine, R., Loayza, N., Beck, T. (2000). Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics* 46:31–77.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics* 30:67–80.

Nizalova, O. Y., Murtazashvili, I. (2016). Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term. *Journal of Econometric Methods* 5:71–77.

Rajan, R. G., Zingales, L. (1998). Financial dependence and growth. *American Economic Review* 88:559–586.

Stock, J. H., Wright, J. H. (2000). GMM with weak identification. *Econometrica* 68:1055–1096.

Stock, J. H., Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Andrews, D. W. K., Stock, J. H. eds., *Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg*, New York: Cambridge University Press, pp. 80–108.