Maastricht University

UNITED NATIONS UNIVERSITY

UNU-MERIT

Working Paper Series

#2022-002

Estimation of a production function with domestic and foreign capital stock

Thomas Ziesemer

Published 10 January 2022

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Boschstraat 24, 6211 AX Maastricht, The Netherlands
Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT

Estimation of a production function with domestic and foreign capital stock

Thomas Ziesemer, Department of Economics, Maastricht University, and UNU-MERIT. Address: P.O. Box 616, NL 6200 MD Maastricht. E-mail: T.Ziesemer@maastrichtuniversity.nl. ORCID: 0000-0002-5571-2238.

Abstract We estimate a Cobb-Douglas production function distinguishing between a domestic and a foreign capital stock built from data of imported machinery and transport equipment for Brazil. The preferred regression uses log levels estimated by GMM-HAC. Results are that the elasticity of production of foreign capital is about 40% of that of domestic capital, the function has constant returns to scale in capital and labour variables, and human capital and technical change are also highly productive. JEL codes: C22, C51, E23, F43, O54. Keywords: time-series, estimation, production function, open economy, Brazil.

1. Introduction

Bardhan and Lewis (1970) have merged the two-gap model of Chenery and Bruno (1962) with the growth model of Solow (1956). The result is a neoclassical growth model with a domestic and a foreign capital stock.[1] The investment in the foreign capital stock has to be paid for, by exports, sooner or later if foreign debt is included as in the extensions under perfect or imperfect capital mobility (Ziesemer 1995, 1998). Compared to the neoclassical closed-economy model the strength of the model is that (i) it deals with growth through capital accumulation linked to international trade in capital goods, consumption goods and foreign debt, the major aspects of globalization, and (ii) it has a steady-state growth rate with income and price elasticities of export demand because world income enters the export function, besides technical change. However, the related production function with domestic and foreign capital stocks has never been presented as an empirical estimate. Such an estimate is the contribution of this paper.

2. The model

The production function that we estimate in this paper is of the Cobb-Douglas type:

$$Y = e^{c+bt+v_t} K_d^{\alpha} e^{\beta H} T_h^{\gamma} [L(1-u)^{\theta} K_f^{\mu}, \tag{1}$$

Y denotes output, K_d is domestic capital, K_f foreign capital, *H* human capital, *L* labour force, *u* the unemployment rate, T_h labour augmenting technical change taking into account human capital in its calculation, *c* a constant, *t* a time trend, and v_t a stochastic term. Entering human capital in the same form as Barro and Lee (2013) do results in having βH in the log-linear regression; this works slightly better here than the log(H) version.

Taking logs of (1) we get

$$logY = c + bt + \alpha logK_d + \beta H + \gamma logT_h + \mu logK_f + v_t \tag{1'}$$

As the variables may have unit roots, we should also consider estimating the function in first differences:

[1] Imported capital goods are also increasingly recognized as important in the heterodox literature (Blecker 2021).

$$dlogY = b + \alpha dlogK_d + \beta dH + \gamma dlogT_h + \mu dlogK_f + v_t - v_{t-1} \qquad (2)$$

In (1') and (2) residuals are different. In (2) we have a special case of a moving average. The stochastic term may have serial correlation with lags *j* and moving average residuals ε_{t-i}.

We want to link them to an ARMA process (see Davidson and MacKinnon 2004, chapter 13)

$$v_t = \sum_{j=1}^{p} \rho_j \, v_{t-j} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \qquad (3)$$

The process with only the first sum on the right-hand side is called an autoregressive process of order j, *ar(p)*, and with only the second sum it is called a moving average process, *ma(q)*. Together they are called an ARMA(p,q) process. ρ_j and θ_i will be found through the estimation.[2] The related parameters have to be found in the estimation together with the exponential parameters, which are elasticities of production. A level model will then combine (1') and (3), and a difference model will combine (2) and (3) and is typically called ARIMAX model, where the 'I' stands for 'integrated' and the X for the regressors other than the constant. Note that if the time trend is statistically insignificant in the level model, then the intercept may be statistically insignificant in the difference model. Using (1') and its lagged form, we insert the residuals into the autoregressive process (3) and estimate the level model

$$logY_t = c + bt + \alpha logK_{d,t} + \beta H_t + \gamma logT_{h,t} + \mu logK_{f,t} + \sum_{j=1}^{p} \rho_j \, (logY_{t-j} - c - bt - \alpha logK_{d,t-j} - \beta H_{t-j} - \gamma logT_{h,t-j} - \mu logK_{f,t-j}) + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \qquad (4)$$

Similarly, using (2) and its lagged form, we insert the residuals into the differenced version of the autoregressive process (3) and estimate the differenced model

$$d(logY_t) = b + \alpha d(logK_{d,t}) + \beta d(H_t) + \gamma d(logT_{h,t}) + \mu d(logK_{f,t}) + \sum_{j=1}^{p} \rho_j \, [d(logY_{t-j}) - b - \alpha d(logK_{d,t-j}) - \beta d(H_{t-j}) - \gamma d(logT_{h,t-j}) - \mu d(logK_{f,t-j})] + d\varepsilon_t + \sum_{i=1}^{q} \theta_i d\varepsilon_{t-i}. \qquad (5)$$

We will present estimates of special cases of these models in which some of the ρ_j, θ_i are zero. The residual in (5) is in differences, which are moving averages, indicating overdifferencing. Overdifferencing is not a problem if the serial correlation is taken into account (Maddala and Kim 1998).

3. The data

We use data for Brazil. We take output as GDP in constant 2010 local currency units, unemployment rates, and labour force data from World Development Indicators (WDI), (World Bank 2021). As unemployment rates have gaps, we run a regression for Okun's law, make a forecast and use its values to fill the gaps. We use technical change data from Ziesemer (2021) selecting the elasticity of substitution close to unity with a CES parameter of 0.99 corresponding almost exactly equal to the Cobb-Douglas function because it is contained there in both selection procedures; these levels have fallen since about 1980 in a similar way as the TFP data from PWT9.1, which are constructed slightly differently. We construct the foreign capital stock from imported machinery and transport equipment in current 1000$ (1989-2020 from the World Integrated Trade Solution, WITS), multiply it by 1000 and the official exchange rate, divide by the GDP deflator from WDI and multiply it by 100. Then we apply the perpetual inventory method to this investment variable. For this we use the average depreciation rate of 4.25% from PWT9.1 for the years 1989-2017, which also enters the

[2] For firm-level panel data Blundell and Bond (2000) use an ar(1) process.

construction of the initial value found as the 1989 imported investment goods divided by the rate of depreciation plus a growth rate of 4.7%, which is the capital growth rate for Brazil in 1989 in Ziesemer (2021). The domestic capital stock is obtained in the same way, based on gross fixed capital formation data from WDI diminished by the imported machinery and transport equipment; by construct, this variable also includes investment in buildings. Human capital data are taken from PWT9.1; they are constructed as index between 1 and 5, and therefore can grow only to its upper limit and act as shifters of the production function rather than permanently growing factors.

4. Econometrics, estimation results and interpretation

In Table 1, column 1, we show a least-squares estimate for the level model (4). The elasticity of production of the foreign capital stock is only 40% of that of the domestic capital stock, here and in the next two regressions in columns 2 and 3. Time trends are never significant because we include a technical change variable. All other variables have elasticities of production as common in the literature. The low elasticity of production may be explained by issues of technology transfer as discussed in the literature on appropriate technology. When considering returns to scale we should not include human capital because its index maximum value of 5 does not allow taking arbitrary multiples. Returns to scale are close to unity with a p(crs) = 0.8 in a Wald test.
All variables are not independent of the GDP because they are end of period values including current investments and are thereby endogenous. Therefore, we use also the two-stage least squares method in the second column and GMM in the third column of Table 1. Least squares and two-stage least squares consider heteroscedasticity and serial correlation only in the standard errors and covariances. GMM estimators take them into account also in the coefficient estimate shown in Table 1, column 3**.** The elasticities of production are now higher for human capital and technical change and slightly lower for domestic capital and labour variables than those of Table 1, column 1 and 2. The result is close to constant returns to scale.

The ar(5) coefficients probably indicate a business cycle effect.[3] As we do not use data for machine and labour hours, factors are fully used in booms, but in other periods there is lower output whereas the lower use of the factors' hours is not captured in our data. In the production function as defined for these data, the economy is below the function, not on it, whenever factors run less hours and this may lead to ar(5) terms indicating that every 5 years the economy is in a similar situation. Moving-average terms make GMM estimates dependent on initial values and thereby unstable. We have dropped them from the analysis of the level equation because removing differencing in the form of ar(1) terms (see Nau 2020) does not solve the instability problem.

Moreover, having a GMM result suggests that the estimate of the covariance matrix has converged. However, with additional observations it might increase if there are unit roots (Davidson and MacKinnon 2004, ch. 14). A vector-error-correction model based on a VAR with lag length one based on the SIC criterion (because of the low numbers of observations) would suggest that the model has five cointegrating equations (r = 5) from the trace test or two (r = 2) from the maximum-eigenvalue test, which would imply one or four (K-r= 6-r) unit eigenvalues in the system of cointegrated equations. For a VECM estimate, we have a too small number of observations. Therefore, we also estimated the differenced model (5), although (i) the models of column 1 to 3 have been tested for ar(1) terms and actually include ar(2) and ar(5) terms and (ii) the values of Durbin-Watson statistic seem to preclude first-order serial correlation as in the presence of unit roots without cointegration. A simple least-squares result in column 4 of Table 1 without any ar or ma terms leads to very

[3] The econometrics literature often discusses this under seasonal effects, which are similar to business cycle effects with a less clear number of periods in contrast to the four seasons.

plausible elasticity and constant-returns-to-scale results but the Durbin-Watson statistic signals high first-order serial correlation. Adding ARMA(p,q) terms (for ma terms see notes to column 5, Table 1) and using instruments against endogeneity again we get the results in column 5 of Table 1 from 2SLS (two-stage-least squares) estimation. The moving averages in the results for a differenced model using GMM (not shown) turn out to be in first differences as in equation (5) and this suggests removing the differencing (Nau 2020), referring us back to the level approach as in columns 1-3 of Table 1. Moreover, for the differenced model estimated with GMM-HAC estimate we cannot avoid the problem of weak instruments (not shown).

Table 1: Estimation results for the level and differenced models

Variable, Method	Least sq.(a)	2SLS (b)	GMM-HAC(c)	Least squares (differenced)(d)	2SLS (differenced) (e)
constant	4.26 (2.33)	3.77, (1.68)	4.46 (3.02)	-0.009, (-1.00)	-0.008 (-2.51)
$\log K_d$	0.2865 (4.7)	0.3, (4.44)	0.283 (6.22)	0.256, (1.834)	0.287 (5.99)
H	0.278 (6.26)	0.264, (4.90)	0.283 (7.56)	0.363, (3.158)	0.35 (6.23)
$\log T_h$	0.697 (16.5)	0.69, (12.4)	0.71 (25.8)	0.6, (9.60)	0.75 (28.25)
Log(L*(1-u))	0.625 (10.7)	0.615, (9.4)	0.61 (16.67)	0.6, (5.83)	0.539 (9.87)
$\log K_f(-2)$	0.115 (3.61)	0.123, (3.94)	0.12 (5.18)	0.16, (1.83)	0.226 (5.16)
AR terms	ρ_2= 0.34 (2.52) ρ_5 = -0.2 (-2.16)	ρ_2 = 0.349, (2.49) ρ_5 = -0.196, (-1.96)	ρ_2 = 0.358 (2.91) ρ_5 = 0.21 (3.85)	-	ρ_1= -0.357 (-1.87)
Adjusted sample	1995-2017	1995-2017	1995-2017	1991 - 2017	1997-2017
Adj.R-sq., J-stat. (p(J))	0.9995, -	0.9995, 11.74 (0.3)	0.9995, 8.31(0.6)	0.92, -	0.9877 12, (0.446)
Durbin-Watson st.	2.16	2.13	2.17	1.41	2.64
Andrews bandwith, (f)	1.1765	1.075	1.1716	3.58	2.8121
returns to scale (g)	1.0265	1.038	1.012	1.02	1.052

Dependent Variable: LOG(Y); d(log(y)) in column 4 and 5. T-values below coefficients in parenthesis. HAC standard errors & covariance (Bartlett kernel). (a) ARMA Conditional Least Squares (Gauss-Newton/Marquardt steps); $p \leq 0.0474$. (b) Instrument specification: C, LOG(KD(-1)), (H(-1)), LOG(TH099(-1)), LOG(L(-1)*(1-U2(-1))), LOG(KF47(-2)); constant insignificant, other $p \leq 0.069$; Lagged dependent variable & regressors from ar(2) and ar(5) terms added to instrument list; no IV dropped. (c) s.e. heteroscedasticity & autocorrelation consistent; weighting matrix: HAC (Bartlett kernel, Andrews bandwidth = 1.23; Instrument specification C, LOG(KD(-3)), (H(-3)), LOG(TH099(-3)), LOG(L(-3)*(1-U2(-3))), LOG(KF47(-2)) is stronger than using lag 1; Sequential 1-step weighting matrix & coefficient iteration; MA Backcast: 1993 1994; $p \leq 0.0086$; lagged dependent variable & regressors from ar(2) and ar(5) terms added to instrument list; no IV dropped. (d) Least Squares; Newey-West HAC standard errors & covariance (Bartlett kernel); constant insignificant, other $p \leq 0.1$. (e) MA Backcast: 1992-1996; MA terms θ_1= -1.18 (-145.9), θ_5 = 0.274 (33.28); instruments: all regressors with lags 1,2,5,6 except d(log(Kf47(-2))) (automatically de-selected). (f) Bartlett kernel. (g) Sum of coefficients of Kd, Kf, labour.

A regression suffering less from weak instruments is the 2SLS regression in column 5 of Table 1. The instruments for technology and labour are weak, but the regression does not depend on initial values. All coefficients are a bit higher here and we find slightly increasing returns to scale, although with a p = 0.397 for the constant returns hypothesis. Another weak point clearly is the Durbin-Watson statistic of 2.64. But the 2SLS regression for the differenced model still shows that in case of unit roots and differencing there is a reasonable estimate. The difficulty here is the choice of valid instruments. Denoting the regressor matrix as *X*, our endogeneity assumption is *E(X'u) >0*, implying *E(X(-l)'u(-l)) >0*. Using lags as instruments imposes *E(X(-l)'u) = 0*. The instrument selection presented in note (e) to Table 1 takes this into account by way of using instruments with lag *q+1* for ma terms with *q = 1*, *5*.

Instead of choosing between level and difference model we can also estimate (4) and (5) as a system of equations, where the corresponding difference and level terms have the same coefficients except for the constants. Table 2 contains the results. The weighted-least-squares method multiplies the equation by estimated inverse conditional variances. The SUR method takes into account the contemporaneous correlations of the residuals of the two equations. The 3SLS method also does so but uses also instruments to deal with endogeneity. THE GMM-HAC method uses instruments and uses heteroscedasticity and serial correlation consistent (HAC) estimation for coefficients and standard errors including contemporaneous correlation of the residuals. By and large, variables have similar size of coefficients across methods in Table 2. The exceptions are the labour coefficient in column 5, which is a bit higher and the foreign capital coefficients in columns 3 and 5, which are a bit lower.

The J-statistic for the IV estimators should be chi-square distributed. It should not be too high and its p-value therefore not too low to be in the chi-square distribution (Davidson and MacKinnon 2004); but it should also not be too low and its p-value too high, because that would mean that instruments do too little (Roodman 2009). These results judge about instruments and specification. All IV estimations in Table 1 have a reasonable p(J) and there are no weak instruments. Only the GMM-HAC estimates take heteroscedasticity and serial correlation into account in a consistent way. A p(J) = 0.598 in column 3 of Table 1 indicates avoiding a too high or too low J-statistic. The GMM-HAC estimate in levels is our preferred regression because instruments are well correlated with regressors (see appendix), estimation is heteroscedasticity and serial correlation consistent (HAC), and the J-statistic is neither too high nor too low. It has almost constant returns to scale[4] and a strong impact of technical change and human capital and thereby dynamically increasing returns to scale. The elasticities of production of capital add up to 0.4, a standard value for aggregate capital in the literature (see Perkins et al. 2013). Other estimators in Table 1 and 2 do not have all of these properties but still show similar results. The system version of GMM-HAC in Table 2 suffers from a low J-statistic and a high p(J) indicating that instruments do too little of bias correction. Nevertheless, the differenced single equation model and the system model support the idea of having a production function with foreign capital.

Modern developments in econometrics have moved from AR(I)MA models to the more general case of ARDL models including error-correction models (see Cho et al. 2021). However, in our case of a low number of observations where (V)ECMs do not work, AR(I)MAX models may be a good way out because they estimate less parameters.

[4] The p-value for the hypothesis of constant returns to scale is 0.8, implying that the crs hypothesis cannot be rejected.

Table 2: Estimation results for the system model

Variable, Method	Iterative LS	Weight. LS (a)	SUR (b)	3SLS (c)	GMM HAC (d)
Constant (differenced eq.)	-0.0023 (-1.09)	-0.0027 (-1.4)	-0.0007 (-0.95)	0.0002, (0.106)	-0.003 (-6.44)
Constant (level eq.)	5.935 (1.67)	4.0 (1.79)	5.267 (2.41)	3.71 (1.84)	3.08 (3.04)
$\log K_d$	0.242 (1.94)	0.293, (4.01)	0.246 (3.61)	0.292, (4.62)	0.25 (8.56)
H	0.274 (4.01)	0.259, (5.14)	0.309 (6.19)	0.26, (5.56)	0.258 (12.22)
$\log T_h$	0.601 (17.29)	0.673, (26.2)	0.693 (27.18)	0.688, (28.76)	0.617 (47.4)
Log(L*(1-u))	0.6 (8.84)	0.613, (11.97)	0.6716 (13.21)	0.634, (12.14)	0.8 (40.2)
$\log K_f(-2)$	0.122 (2.27)	0.128, (3.95)	0.089 (3.05)	0.124, (4.26)	0.089 (5.54)
AR terms diff. eq.	-	ρ_1 =- 0.333 (1.85)	$\rho_1 = -0.333$ (-2.06) ρ_5 =-0.25 (-1.85)	ρ_2 =0.4 (2.976)	$\rho_1 = 0.249$ (4.16)
AR terms level eq.	ρ_1 =0.856 (7.64)	ρ_2 =0.377 (3.55) ρ_5 =-0.176 (1.7)	ρ_2 =0.31 (3.83) ρ_5 =-0.134 (-1.91)	ρ_2 =0.375 (4. 31) ρ_5 =-0.17 (-2.22)	ρ_1 =0.769 (29.16)
Adjusted sample	1991-2017	1992-2017	1995-2017	1995-2017	1992-2017
Det. resid. covariance	2.39E-10	5.11E-10	1.24E-10	2.06E-10	5.35E-10
Adj.R-sq.: diff eq. Level eq.	0.92, 0.999	0.913, 0.9995	0.93, 0.9995	0.943, 0.9995	0.898, 0.99925
Durbin-Watson:diff eq Level eq	1.41, 1.45	1.985, 1.955	1.75, 1.93	2.31, 2.11	2.03 2.32
returns to scale (g)	0.965	1.034	1.007	1.05	1.052

(a), (b), (d): Iterate coefficients after one-step weighting matrix. (c) Sequential weighting matrix & coefficient iteration. Instruments: C, D(LOG(KD(-4))), D(H(-4)), D(LOG(TH099(-4))), D(LOG(L(-4)*(1-U2(-4)))), D(LOG(KF47(-2))) for differenced equation C, LOG(KD(-1)), (H(-1)), LOG(TH099(-1)), LOG(L(-1)*(1-U2(-1))), LOG(KF47(-2)) for level equation, and lagged dependent variables and regressors from ar terms added to instrument list of level and difference equation; lag 4 instruments are stronger for difference equation than lag 3. (d) Kernel: Bartlett, Bandwidth: Variable Newey-West (11), No prewhitening; J-statistic 0.116, p(J)=1; Instruments: C, D(LOG(KD(-2))), D(H(-2)), D(LOG(TH099(-2))), D(LOG(L(-2)*(1-U2(-2)))), D(LOG(KF47(-2))) for the difference eq., C, LOG(KD(-2)), (H(-2)), LOG(TH099(-2)), LOG(L(-2)*(1-U2(-2))), LOG(KF47(-2)), and lagged dependent variables and regressors from ar terms added to instrument list of level and difference equation.

5. Conclusion and suggestions for further research

Our estimates show that a production function with foreign capital as used by models based on Bardhan and Lewis (1970) can be defended as empirically realistic and therefore the models have a solid empirical basis. However, the autoregressive processes employed here are under suspicion of hiding mis-specification in some parts of the literature. This would be the case if actually the production functions are of a more general CES or VES type. The elasticities of production, which are now constant coefficients, then would be functions of the output/factor ratios and marginal products and would vary over time if the economy is not in a steady state. Future research may try to estimate a more general CES function or a VES function distinguishing between domestic and foreign capital.

References

Bardhan, P.K., Lewis, S. (1970). Models of growth with imported inputs. Economica, 37 (148), 373-385.

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. Journal of development economics, 104, 184-198.

Blecker, Robert A. (2021) New advances and controversies in the framework of balance-of-payments-constrained growth. Journal of Economic Surveys, 1–39.

Blundell, Richard & Stephen Bond (2000) GMM Estimation with persistent panel data: an application to production functions, Econometric Reviews, 19:3, 321-340.

Chenery, HB, M Bruno (1962). Development alternatives in an open economy: the case of Israel. The Economic Journal 72 (285), 79-103.

Cho, Jin Seo, Matthew Greenwood-Nimmo, Yongcheol Shin (2021) Recent developments of the autoregressive distributed lag modelling framework. Journal of Economic Surveys, 1–26.

Davidson, R. and J.G. MacKinnon (2004) Econometric Theory and Methods. Oxford University Press. New York et al.

Keane, Michael and Timothy Neal (2021) A Practical Guide to Weak Instruments. Mimeo.

Maddala, G.S. and I.-M. Kim, 1998, Unit Roots, Cointegration and Structural Change, Cambridge University Press.

Nau, Robert (2020) Statistical forecasting: notes on regression and time series analysis. https://people.duke.edu/~rnau/411home.htm.

Perkins, D.H, S. Radelet, D.L. Lindauer and S.A. Block (2013) Economics of Development, 7th ed., Norton.

Roodman, D. (2009) A note on the theme of too many instruments. Oxford Bulletin of Economics and Statistics 71 (1), 135–158.

Solow, R.M. (1956) A Contribution to the Theory of Economic Growth, Quarterly Journal of Economics, Vol. 70, pp. 65-94.

World Bank (2021) World Development Indicators (WDI). Washington.

Ziesemer, Thomas (1995) Growth with Imported Capital Goods, Limited Export Demand and Foreign Debt. Journal of Macroeconomics 17(1), 31-53.

Ziesemer, Thomas (1998) A Prebisch-Singer Growth Model and the Debt Crises, in: "DEVELOPMENT ECONOMICS and POLICY", edited by David Sapsford and John-ren Chen, Macmillan, 300-317.

Ziesemer, Thomas (2021) Labour-augmenting technical change data for alternative elasticities of substitution: growth, slowdown, and distribution dynamics, Economics of Innovation and New Technology, DOI: 10.1080/10438599.2021.1956316. Online.

Appendix

Regression of regressors on instruments for the preferred GMM HAC regression in column 3 of Table 1, not using regressors that serve as their own instruments. Estimation Method: Weighted Least Squares. P-values in parentheses.

GMM-HAC in levels (Table 1, column 3)

LOG(KD) = 0.25995 + 0.99*LOG(KD(-3)), R-sq. = 0.966, Durbin-Watson stat = 0.21, Observations: 30,
(0.79) (0.00)

H = -0.177 + 1.14*H(-3), R-sq. = 0.997, Durbin-Watson stat = 0.219, Observations: 57
(0.00) (0.00)

LOG(TH099) = 0.3 + 0.77*LOG(TH099(-3)), R-sq. =0.633, Durbin-Watson stat = 0.369, Observations: 55
(0.0042) (0.00)

LOG(L*(1-U2)) = 2.7 + 0.85*LOG(L(-3)*(1-U2(-3))), R-sq. = 0.96, Durbin-Watson stat = 1.03, Observations: 28
(0.00) (0.00)

Lag 3 should be far enough in the past to suggest exogeneity. The regressions suggest strong instruments because the coefficients, their significance and the non-adjusted R-squared are high (added for readers familiar with the weak instruments literature). For the level regression on the third lag, the correlation for domestic capital and human capital with coefficients near unity and high R-squared may be "too good", implying that the instrument is almost perfectly correlated with the regressor and therefore also with the residual. However, the Durbin-Watson statistics here are 0.21 and 0.22.

The related first-stage regressions of the regressors on *all* instruments have adj $R^2 = 0.999$ and $p(F) < 0.013$ under the assumption of homoscedasticity. The F-statistics are clearer higher than the traditional requirement of $F > 10$ for the case of homoscedasticity. However, we clearly have heteroscedasticity.

Table A.1: F-statistics of first-stage regression related to GMM HAC estimate (a)

Regressor as dependent variable of first-stage regression	homoscedasticity	Heteroscedasticity (b)	
		Andrews kernel bandwidth	Newey-West HAC kernel bandwidth
logKd	1407.95 (0.000000)	12777.98 (0.000000)	122829.6 (0.000000)
H	2050.9 (0.000000)	14780.05 (0.000000)	110156.1 (0.000000)
log(th099)	22.9 (0.0013)	912.8 (0.000000)	9085.759 (0.000000)
L(1-u)	285.35 (0.000000)	4213.7 (0.000000)	17880.27 (0.000000)

(a) See Table 1, column 3; p-values in parentheses. (b) Coefficient covariance (HAC Newey West), Bartlett kernel.

Wald F-statistics for the case of heteroscedasticity are much higher than those for homoscedasticity, especially if the Newey-West-automatic-bandwidth procedure for the standard errors is used in Table A.1. However, the exact requirements for both homoscedasticity and heteroscedasticity are still in the econometric discussion (Keane and Neal, 2021).

The UNU-MERIT WORKING Paper Series