CFDS Discussion Paper Series

State-Level Capital and Investment: Refinements and Update

Makram El-Shagi, Steven Yamarik

Discussion Paper 2018/1

AUTHORS

Makram El-Shagi

HenU Center for Financial Development and Stability
Henan University
E-mail: makram.el-shagi@cfds.henucon.education
Tel: +86 155 6511 5281

Steven Yamarik (Corresponding Author)

Department of Economics
California State University Long Beach, CA
E-Mail: Steve.Yamarik@csulb.edu

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of the CFDS. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors. Comments and suggestions on the methods and results presented are welcome.

IMPRESSUM

© CFDS, 2018

HenU Center for Financial Development and Stability
Dongliuzhai Building, 85 Minglun Street
Henan University, Minglun Campus
Shunhe, Kaifeng, Henan, China
Tel. +86 (30) 897 89-0
<http://cfds.henucon.education>

Papers can be downloaded free of charge from the CFDS website:
<http://cfds.henucon.education/index.php/research>

State-Level Capital and Investment: Refinements and Update

Makram El-Shagi^a and Steven Yamarik^{*b}

^aSchool of Economics, Henan University, Kaifeng, Henan, China,
475001

^bDepartment of Economics, California State University, Long
Beach, CA, USA, 90840.

February 28, 2018

Abstract

This paper presents updated estimates for state-level capital and investment for 1950 to 2015. We improve upon the procedure of Garofalo and Yamarik (2002) and Yamarik (2013) by using quantity measures to apportion the mining capital stock and a geometric pattern of depreciation to derive investment data. In an empirical application we use our data to estimate the production function and a simple growth model. We find coefficient estimates that support constant returns to scale and a 1/3 output elasticity of capital. These results are consistent with past regional and cross-country papers, supporting the plausibility of our data.

JEL Classification: O47, O51, R11

Keywords: capital stock, investment, production function, Solow

^{*}Corresponding author, email: Steve.Yamarik@csulb.edu

1 Introduction

U.S. states provide a ready-made setting to test macroeconomic theories of production, growth and factor use because they possess similar institutional backgrounds which reduces potential omitted variable bias. One problem confronting state-level analysis is the lack of an up-to-date series on physical capital and investment. The Bureau of Economic Analysis (BEA) provides estimates for the net capital stock for the U.S., but not for individual states. Garofalo and Yamarik (2002) and Yamarik (2013) filled this data gap by constructing state-level capital stock and investment data for 1947-2007 and 1990-2010. We improve upon their methodology and update their data up to 2015.¹ By providing more accurate estimates of state-level capital and investment, we provide an essential asset for the large empirical literature that analyzes production (Brock and Ogloblin, 2014; Benos et al., 2016); total factor productivity (Sharma et al., 2007; Cardarelli and Lusinyan, 2015); employment (Liu, 2013); foreign direct investment (Mullen and Williams, 2005; Bode and Nunnenkamp, 2011); tax policy (Reed, 2008); public capital (Kalyvitis and Vella, 2015); growth (Reed, 2009; Turner et al., 2013); human capital (Yamarik, 2008, 2011); inequality (Beenstock et al., 2011); immigration and migration (Peri, 2012; DiCecio and Gascon, 2010); resource use (Metcalf, 2008; Burnett and Madariaga, 2017); and income convergence (Magrini, 2004; Heckelman, 2013).

We improve upon the earlier procedure of Garofalo and Yamarik (2002) in two important ways. First, we divide the mining sector into its 3-digit sub-sectors and then use quantity measures as weights for the volatile energy sub-sectors. Derived from production data, these quantity weights are more closely linked to capital-use than relative earnings which are influenced by relative price fluctuations. Second, we use a geometric pattern of depreciation provided by the BEA to derive our investment data. The geometric pattern generates higher (and more plausible) rates of investment and depreciation than the linear pattern used previously.

We assess the plausibility of our data by estimating a Cobb-Douglas production function and Solow growth model. Each model provides testable predictions on the output elasticities of capital and investment. We find empirical support for each prediction using a host of widely-used estimator.

The remainder of the paper is organized as follows. Section 2 discusses the data methodology and the adjustment to the mining sector. Section 3 presents a brief overview of our updated data estimates. Section 4 discusses the estimation of a Cobb-Douglas production function, while Section 5 examines the Solow growth model. Section 6 shows the results of each theoretical prediction in graphical form.² Section 7 concludes.

¹The data is available for download at jg and will be updated annually at the website.

²Detailed results in table form can be found in the Appendix.

2 Data methodology

2.1 Procedure

We start with the original procedure of Garofalo and Yamarik (2002) and Yamarik (2013) to create state-by-state capital stock and investment estimates for 1950-2015. The procedure apportions the national capital stock estimates to the states using either SIC or NAICS 2-digit industry-level earnings data.³

For each 2-digit industry, the national capital stock is divided up based on the relative earnings generated within each state:

$$k_{i,j}(t) = \left[\frac{earn_{i,j}(t)}{EARN_i(t)} \right] K_i(t) \quad (1)$$

where the subscripts i represents the 2-digit industry, j denotes the state, and t represents the year. The lowercase $earn_{i,j}$ refers to state-level earnings in industry i , while the uppercase $EARN_i$ and K_i denote national earnings and capital in industry j .

The capital stock for each state is then the sum of the 2-digit estimates:

$$k_j(t) = \sum_{i=1}^{10/19} k_{i,j}(t) \quad (2)$$

where i represents the 10 SIC or 19 NAICS 2-digit industries.

We use the perpetual inventory method to generate gross investment and depreciation rates for U.S. states. The BEA applies a geometric pattern of depreciation to capital equipment to produce national depreciation values DEP_i for each industry i . We use lagged earnings weights to allot depreciation to each state and then sum across industries:

$$dep_{i,j}(t) = \left[\frac{earn_{i,j}(t-1)}{EARN_i(t-1)} \right] DEP_i(t) \quad (3)$$

³The Standard Industry Classification (SIC) was the classification system used until replaced by the North American Income Classification System (NAICS) in 1997. The SIC is a 5-digit hierarchical system, offering four levels of detail. There are **ten** 2-digit SIC private industries: Farms (81), Agricultural Services (100), Mining (200), Construction (300), Manufacturing (400), Transportation and public utilities (500), Wholesale trade (610), Retail trade (620), Finance, insurance, and real estate (700), Services (800). The NAICS is a 6-digit hierarchical system, offering five levels of detail. There are **nineteen** 2-digit NAICS private industries (with code): Agriculture, Forestry, Fishing and Hunting (11); Mining, Quarrying, and Oil and Gas Extraction (21); Utilities (22); Construction (23); Manufacturing (31-33); Wholesale Trade (42); Retail Trade (44-45); Transportation and Warehousing (48-49); Information (51); Finance and Insurance (52); Real Estate and Rental and Leasing (53); Professional, Scientific, and Technical Services (54); Management of Companies and Enterprises (55); Administrative and Support and Waste Management and Remediation Services (56); Educational Services (61); Health Care and Social Assistance (62); Arts, Entertainment, and Recreation (71); Accommodation and Food Services (72); Other Services, except Public Administration (81).

$$dep_j(t) = \sum_{i=1}^{10/19} dep_{i,j}(t) \quad (4)$$

where i represents the 10 SIC or 19 NAICS 2-digit industries.

Using the capital stock (2) and depreciation value (4), a gross investment series is then estimated for each state j :

$$inv_j(t) = k_j(t) - k_{j-1}(t) + dep_j(t). \quad (5)$$

We divide the depreciation value by the lagged capital stock to generate a depreciation rate for each state:

$$\delta_j(t) = dep_j(t)/k_j(t-1). \quad (6)$$

Our data methodology assumes that the output-to-capital ratio within each industry is constant across U.S. states. This assumption implies that each industry is in a common steady state. For example, the manufacturing industry in Pennsylvania is assumed to be in the same steady state as its counterpart in Ohio. What if the manufacturing industry is converging towards its steady state? The output-to-capital ratio would be below its steady-state value. This only poses a problem for the methodology if the initial output-to-capital ratios differ across U.S. states. If the ratios do vary, then (1) and (2) would apportion too much to those states farther from steady-state and too little to those states closer to steady state. What if the manufacturing industry is in different steady states with different capital-to-output ratios? If so, then the data methodology would apportion too much to those states with lower ratios and too little to those state with higher ratios. However, this possibility is unlikely given that competition would lead firms in all states to adopt the best production technology available.

2.2 Imputation of Missing Values

One problem arising from both the SIC and NAICS data is the occasional omission of industry-level earnings data to avoid disclosure of confidential data. These omissions are quite rare, occurring in the Agriculture (2.7%); Mining (3.4%), Utilities (0.8%), Construction (0.4%), Transportation (0.4%), Information (0.4%) and Other Services (0.4%). We use a linear interpolation for a missing value when data is available before and after. If the omission is current, we extend the latest value available forward. However, the specifics of our imputation have little to no impact on the state-level capital stock estimates due to the minuscule contribution of the omitted states to the industry.

2.3 Adjustment to the Mining Industry

The Mining, Quarrying, and Oil and Gas Extraction (21) industry poses the biggest challenge to our methodology. Under the old SIC system, mining (21) encompasses four 3-digit industries - Metal Mining, Bituminous Coal and Lignite Mining, Oil and Gas Extractions, and Mining and Quarrying of Nonmetallic Minerals except of fuels. Under the current NAICS system, mining is comprised of three 3-digit NAICS sub-sectors - Oil and Gas Extractions (211), Mining except Oil and Gas (212), and Support Activities for Mining (213). In terms of capital, the Oil and Gas Extractions (211) is the most important sub-sector, comprising around 80 - 85% of earnings under both classification systems.

The prices of oil, natural gas and coal are subject to large and variable price shocks (Jacks et al., 2011). As a result, the earnings in these sub-sectors are subject to the same volatility. With states possessing different mixes of these sub-sectors; the price shocks can lead to changes in state-level mining earnings that diverge from the underlying production mix.

We address this issue by dividing the mining industry into its 3-digit sub-sectors and then use quantity measures for weights in the volatile energy sub-sectors. There are national capital stock data available for each 2-digit sector and 3-digit sub-sector. In the Oil and Gas Extractions sub-sector, we add up oil and natural gas production by converting each into common British Thermal Units (BTU) output. We then use these BTU output units as weights to apportion the national oil and gas capital stock.

$$k_{oilgas,21,j}(t) = \left[\frac{btu_{oil,21,j}(t) + btu_{gas,21,j}(t)}{BTU_{oil,21}(t) + BTU_{gas,21}(t)} \right] K_{oilgas,21}(t) \quad (7)$$

where btu is state-level output and BTU is national output. The subscripts ($oil, gas, oilgas$) represent each 3-digit sub-sector, '21' refers to the 2-digit mining industry, and subscript j denotes the state.

In the Coal Mining sub-sector under the SIC system, we use coal production in BTU output units as weights to apportion the national coal capital

$$k_{coal,21,j}(t) = \left[\frac{btu_{coal,21,j}(t)}{BTU_{coal,21}(t)} \right] K_{coal,21}(t) \quad (8)$$

where the subscripts $coal$ represents the 3-digit sub-sector.

For the other sub-sectors, there are no common output units available so we use earnings at the 3-digit level as weights:

$$k_{d,21,j}(t) = \left[\frac{earn_{d,21,j}(t)}{EARN_{d,21}(t)} \right] K_{d,21}(t) \quad (9)$$

where the subscript d refers to the 3-digit sub-sector NAICS Mining except Oil and Gas (212) or Support Activities for Mining (213).

Figure 1: Comparison of Mining Capital Stock for Top-15 States

We then sum up the 3-digit mining capital estimates to derive a mining capital stock for each state:

$$k_{mine,j}(t) = \sum_{d=1}^{4/3} k_{d,mine,j}(t) \quad (10)$$

where d represents the four SIC or three NAICS 3-digit mining sub-sectors.

Figure 1 presents the mining capital stock estimates using the earnings and output methods for the top-15 states in 2015. These 15 states possess 91% of the national mining capital stock. The main impact of the output adjustment method is to redistribute mining capital away from the top-5 states to the next-7 states. Under the earnings method of equation (1), the top-5 states have 75% of the national mining capital, while the next-10 states have 16%. Under the output adjustment of equations (7)-(10), the capital stock of the top-5 falls to 49%, while the mining capital of the next-10 states rises to 42%. There are minor differences in the other 36 states although they are quite insignificant for their impact on state capital.

The mining capital stock are intermediate estimates in the creation of our state-level capital stock data (2). They are not meant to be used directly in empirical estimation of mining output or other analysis. Nevertheless, we can evaluate their plausibility by looking at the capital-output ratios in the mining sector. The annual **unweighted** average of $(k/y)_{mine}$ ranges from 3.6 to 5.0 under the SIC system and 4.3 to 5.3 under the NAICS system, which are within

the historical national range of 2.0 to 8.0. Although there are a few extreme values, the majority of the $(k/y)_{mine}$ distribution falls inside the national range.

2.4 Combining the SIC and NAICS Series

Our procedure generates one set of estimates for 1947-2000 using SIC earnings data and another set for 1998-2015 using NAICS data. We merge the capital stock and depreciation series in 2000 where the correlation coefficients are 0.98 and 0.99 in log-levels and 0.83 and 0.74 in log-differences. We then convert each series into 2009 dollars using the national capital stock deflator. We then use equation (5) to create a common investment series in 2009 dollars and (6) to generate a common depreciation rate.

3 Overview of data estimates

3.1 Net Capital Stock

Figure 2 shows the distribution of the national capital stock across the eight BEA geographic regions from 1950 to 2015. In 1950, New England, Mideast and Great Lakes accounted for 51%; the Southeast, Southwest and Far West contributed 38%; while Plains and Rocky Mountain possessed the remaining 11%. Over time, there has been a shift away from the Northeast and Midwest to the faster growing South and West. By 2015, the distribution has virtually flipped with New England, Mideast and Great Lakes falling to 36%; the Southeast, Southwest and Far West rising to 53%; and Plains and Rocky Mountains remaining at 11%.

Table 1 presents state-level estimates of the capital stock in 1950, 2000 and 2015. There is a wide dispersion of capital across the states. In 2015, the 10 largest states - California, Texas, New York, Pennsylvania, Florida, Illinois, Ohio, New Jersey, Georgia, and Michigan - account for 56% of the national capital stock, while the 10 smallest states - District of Columbia, New Hampshire, Idaho, Hawaii, Maine, Montana, Delaware, Rhode Island, South Dakota, and Vermont - possess only 3%. Although the states in the top-10 and bottom-10 have changed some, the top-10 states account for 55 to 59% of the national capital stock and the bottom-10 contribute only 3% over time.

Table 1 also shows the annual growth rates of capital for 1950-2015 and for the two subperiods 1950-2000 and 2000-2015. In our state-level data, there is a good deal of variation across states and across time. For the 1950-2015 period, the national capital stock grew at an annual rate of 2.76%, but 34 of the 51 states grew faster than the national rate. In particular, the capital stocks of Alaska, Arizona, Florida and Nevada grew above 5.00 percent per annum, while West Virginia grew less than 1.50%.

There has been a marked slowdown in the growth rate of capital. The growth rate for the national capital stock fell from 3.05 in 1950-2000 to 1.77% in 2000-2015. Likewise, 46 of the 51 states experienced a decrease in the rate

Table 1: Capital stock estimates for select years

State	Capital Stock (millions of 2009 \$)			Growth rates (%)		
	1950	2000	2015	50-15	50-00	00-15
Alabama	44,303	189,644	219,804	2.91	3.64	0.98
Alaska*	8,679	260,979	166,096	5.37	8.51	-3.01
Arizona	16,041	203,765	305,319	5.36	6.35	2.70
Arkansas	29,230	112,531	160,468	3.10	3.37	2.37
California	290,803	2,016,188	2,791,523	4.11	4.84	2.17
Colorado	29,471	294,573	381,312	4.65	5.76	1.72
Connecticut	42,530	236,856	321,662	3.68	4.29	2.04
Delaware	8,777	53,520	54,585	3.32	4.52	0.13
District of Columbia*	20,440	60,047	82,551	2.54	2.69	2.12
Florida	49,819	629,848	847,106	5.15	6.34	1.98
Georgia	46,580	421,274	529,945	4.42	5.51	1.53
Hawaii*	13,284	52,645	70,403	3.03	3.44	1.94
Idaho	12,214	49,739	76,309	3.33	3.51	2.85
Illinois	232,264	705,553	818,746	2.29	2.78	0.99
Indiana	96,068	305,736	398,444	2.59	2.89	1.77
Iowa	48,516	131,642	180,688	2.39	2.50	2.11
Kansas	43,665	153,016	187,219	2.65	3.14	1.34
Kentucky	46,874	167,186	202,680	2.66	3.18	1.28
Louisiana	62,270	246,919	280,621	2.74	3.44	0.85
Maine	17,084	51,617	59,355	2.26	2.76	0.93
Maryland	50,340	227,629	321,244	3.37	3.77	2.30
Massachusetts	92,098	408,357	506,943	3.10	3.72	1.44
Michigan	150,458	534,984	524,986	2.27	3.17	-0.13
Minnesota	57,454	263,588	367,232	3.37	3.81	2.21
Mississippi	23,119	101,703	117,924	2.96	3.70	0.99
Missouri	79,663	255,922	314,189	2.49	2.92	1.37
Montana	16,680	37,684	57,490	2.25	2.04	2.82
Nebraska	22,240	75,538	117,257	3.02	3.06	2.93
Nevada	5,513	99,871	131,688	5.77	7.24	1.84
New Hampshire	9,970	66,853	80,366	3.79	4.76	1.23
New Jersey	98,546	491,371	578,081	3.22	4.02	1.08
New Mexico	19,242	142,680	132,674	3.51	5.01	-0.48
New York	335,039	1,142,358	1,438,622	2.65	3.07	1.54
North Carolina	48,979	349,952	464,737	4.09	4.92	1.89
North Dakota	13,121	38,419	147,785	4.40	2.69	8.98
Ohio	187,185	532,391	651,888	2.27	2.61	1.35
Oklahoma	51,676	207,862	322,916	3.33	3.48	2.94
Oregon	38,557	166,659	216,511	3.14	3.66	1.74
Pennsylvania	263,243	609,054	938,948	2.31	2.10	2.89
Rhode Island	15,098	43,887	53,493	2.30	2.67	1.32
South Carolina	24,738	159,288	215,997	3.94	4.66	2.03
South Dakota	9,332	28,086	49,488	3.03	2.75	3.78
Tennessee	44,864	260,998	324,299	3.60	4.40	1.45
Texas	202,853	1,386,911	1,973,978	4.14	4.81	2.35
Utah	16,618	104,867	168,199	4.21	4.61	3.15
Vermont	8,380	27,860	34,166	2.56	3.00	1.36
Virginia	54,593	332,892	397,745	3.61	4.52	1.19
Washington	58,589	327,429	475,091	3.81	4.30	2.48
West Virginia	61,466	77,574	117,413	1.18	0.58	2.76
Wisconsin	70,813	270,847	349,264	2.90	3.35	1.70
Wyoming	13,991	94,109	114,211	3.82	4.77	1.29
United States	3,303,369	15,210,898	19,839,660	2.76	3.05	1.77

The * states are not included in the production function and Solow growth model analysis.

Figure 2: Distribution of the Net Capital Stock across the 8 BEA Regions

of net capital accumulation. The largest decreases were experienced by Alaska (11.52%), New Mexico (5.49%) and Nevada (5.40%). Only 5 states - Montana, North Dakota, Pennsylvania, South Dakota and West Virginia - experience increases in capital growth from 1950-2000 to 2000-2015.

3.2 Capital-to-labor ratios

Figure 3 shows the capital-to-labor K/L ratios in 2015 divided into five colored categories. The K/L ratios range from a low of 116 in Maine to a high of 646 in Alaska, with a national average of 176. The distribution is skewed to the left with 38 of 51 states possessing K/L ratios below the average. The states with high capital-to-labor ratios have large mining sectors. In particular, Oklahoma (245), North Dakota (397), Wyoming (522) and Alaska (646) have the highest K/L and are the most mining-intensive.⁴

Figure 4 shows the capital-to-labor K/L ratios in non-mining industries in 2015. With mining removed, no state has a K/L in the top category of 230+ and only two states - California (203) and Connecticut (224) - in the second category of 200-230. The K/L ratios of the mining-intensive states Oklahoma (170), North Dakota (178), Wyoming (181) and Alaska (199) fall close to the national average of 176.

⁴The mining sector represents 15 to 25% of state GDP in each of these states relative to 2.5% for the U.S. as a whole.

3.3 Comparison of depreciation and investment rates

Table 2 provides comparison of the depreciation and investment rates produced by Garofalo and Yamarik (2002) (GY) vs. our current estimates (YES) for the overlapping years of 1950-2000. The depreciation rate δ is defined in equation (6). The investment rate i is defined as gross investment divided by personal income. The geometric pattern of depreciation used in YES generates a significantly higher depreciation rate than the linear depreciation pattern used in GY. The 7.0% average depreciation rate in YES is slightly higher than the national depreciation rate of 5.0 to 6.0% (Nadiri and Prucha, 1996; Feenstra et al., 2015). As a result of greater depreciation, the average investment rate nearly doubles from 7.0% in GY to 15.1% in YES. A 15% average investment rate is close to the national fixed investment share of 17% (BEA, 2018).

Table 2: Comparison of investment and depreciation rates

Variable	former GY estimates		current YES estimates	
	Mean	St. Dev.	Mean	St. Dev.
i rate	0.07072	0.03156	0.15138	0.07205
δ rate	0.04474	0.00365	0.07197	0.01285

Note: There are 48 states for 1950-2000 for a total of 2448 observations.

4 Production function estimation

We examine the plausibility of our capital stock data by estimating a two-factor Cobb-Douglas production function: $Y_{i,t} = A_t K_{i,t}^\alpha L_{i,t}^\beta$. Taking logs of each side, we arrive at the unrestricted model:

$$\ln(Y_{i,t}) = a_t + \alpha \ln(K_{i,t}) + \beta \ln(L_{i,t}) + u_{i,t}, \quad (11)$$

where a_t are time fixed effects meant to capture both technological progress and the joint US business cycle and $u_{i,t}$ is an error term. The coefficients α and β are the output elasticities of capital and labor. Most of our models allow for state-specific components in the error term, thus implicitly assuming a more complex production function where highly persistent state-level provisions affect the production process.

We also estimate a restricted model that imposes constant returns to scale: $\alpha + \beta = 1$. As standard in the literature, we estimate a production function where output per worker y depends upon capital per worker k . Neoclassical theory predicts that the elasticity of capital should equal the income share paid to capital. In the national GDP data, the income shares of labor and capital average two-thirds and one-third during the post-war period.

We do not limit ourselves to a single estimator, but instead employ several econometric techniques used in the production function literature. These techniques are static panel (FE, RE, FD), dynamic general method of moments

(GMM), panel cointegration (MG, PMG, DFE), and stochastic frontier analysis (SFA). Our main purpose here is to test the empirical predictions across a range of estimators with different econometric assumptions. To further that purpose, we perform a model averaging process to examine the distribution of the different estimates around our predicted value(s).

4.1 Static panel models

The static panel representation of (11) is

$$\ln(Y_{i,t}) = \alpha \ln(K_{i,t}) + \beta \ln(L_{i,t}) + a_t + v_i + \varepsilon_{i,t}, \quad (12)$$

where the state-specific terms v_i control for unobserved productivity differences across states. The unobserved v_i are typically controlled by using either fixed (FE) or random effects (RE) or by taking first-differences (FD) of the data. Although more efficient, RE is only consistent if the state effects v_i and covariates are independent, which is unlikely in our panel of states. If the data are non-stationary, FE and RE will suffer from spurious regression so the FD estimator is preferred. At the same time, FD will exacerbate any measurement error in the covariates and bias those estimates towards zero (Griliches and Hausman, 1986).

4.2 Panel GMM

One problem with the static panel models is that they cannot address endogeneity of capital and labor. We therefore use a difference GMM estimator of Arellano and Bond (1991) and a system GMM estimator of Arellano and Bover (1995) and Blundell and Bond (1998). The difference GMM estimator uses lagged levels to instrument for the equation in first-differences (following the seminal approach of (Anderson and Hsiao, 1982)). The system GMM estimator adds a levels equation to the first-difference equations and uses lagged differences along with lagged first-differences as instruments. The system GMM estimator improves upon the difference GMM when the data are highly persistent. Both GMM estimators have been applied to the estimation of production functions by Blundell and Bond (2000) and many papers to follow.

Nevertheless there are potential problems in applying system GMM to our data. First, our time dimension is fairly long for a GMM estimator that has been designed for small T and large N samples. In this situation full GMM leads to serious over-instrumentation (Okui, 2009; Roodman, 2009). We therefore present three instrument specifications for each GMM estimator: (i) all available lagged instruments, (ii) a single 2nd lag, and (iii) “collapsing” the instruments into only one instrument per state and lag distance, rather than using one instrument per state, lag distance and year Roodman (2006).

Second, and more difficult to deal with, the stationarity assumptions in system GMM have been hugely underappreciated in the applied literature (Moral-Benito, 2013). While system GMM is well-suited to persistent but stationary

data, its performance suffers greatly if the data actually exhibits a unit root as it does in our application. We present a series of estimators below that deal with these stationarity concerns.

4.3 Dynamic Mean Group

The panel cointegration framework proposed by Pesaran and Smith (1995) and Pesaran et al. (1999) can address both stationarity and endogeneity issues. This framework extends the time-series analysis of Pesaran et al. (2001) to panel data with the following form:

$$\begin{aligned} \Delta y_{i,t} = & \alpha_{i,0}y_{i,t-1} + \alpha_{i,1}k_{i,t-1} + \alpha_{i,2}l_{i,t-1} + \sum_{\tau=1}^p \psi_i \Delta y_{i,t-\tau} \\ & + \sum_{\tau=0}^p \phi_i \Delta k_{i,t-\tau} + \sum_{\tau=0}^p \kappa_i \Delta l_{i,t-\tau} + a_t + v_i + \varepsilon_{i,t}. \end{aligned} \quad (13)$$

The α coefficients for the levels terms capture the long-run relationship of interest, while the differenced terms record the short-run dynamics. We use three different estimators for (13): the mean group (MG) where both short and long run parameters can vary between states (and the average is reported), the pooled mean group (PMG) where only the short-run coefficients are allowed to vary, and the dynamic fixed effects (DFE) where all coefficients (except the state specific constant term) are held constant across states. Pesaran et al. (1999) shows that the PMG estimator has the same asymptotic properties regardless of the stationarity of the regressors.

4.4 Stochastic frontier analysis

Our last estimator is stochastic frontier analysis (SFA) of Aigner et al. (1977). SFA is premised on the idea that deviations of the production measure from the underlying production function have two sources: measurement error and inefficiency. While the former can create deviations in both directions, the latter is a source of strictly negative disturbances because inefficiency reduces production. We apply the original version proposed by Battese and Coelli (1995):

$$y_{i,t} = \alpha k_{i,t} + \beta l_{i,t} + a_t + v_{i,t} - u_{i,t}, \quad (14)$$

where $v_{i,t} \sim \mathcal{N}(0, \sigma_u^2)$ is the measurement error term and $u_{i,t} = e^{-\eta(t-T)} w_i$ with $w_i \sim \mathcal{N}^+(\mu, \sigma_z^2)$ is a technical inefficiency error term. The technical inefficiency error is a product of a time-specific ‘decay’ term $e^{-\eta(t-T)}$ and a state-specific effect w_i .

While stochastic frontier analysis is widely used for micro firm-level data, SFA should only be considered a robustness test in our case. While inefficiency at a firm level basis is a natural occurrence, it is less clear why a state should exhibit the same feature. Even if there is such a thing as aggregate inefficiency,

it would no longer be truncated normal but approaches normality as the number of firms increases.

5 Solow growth model estimation

5.1 Panel estimation and basic model

We next examine the plausibility of our investment and depreciation data by estimating a Solow growth model. The Solow (1956) growth model combines a Cobb-Douglas production function with a capital accumulation equation and exogenous investment, depreciation and labor growth rates. The model is then solved for a steady-state level y^* which depend positively upon the investment rate and negatively upon the depreciation rate. During the transition, convergence predicts that the growth rate should depend negatively upon initial income. In a widely-cited paper, Mankiw et al. (1992) develop an empirical specification of the transitional growth rate:

$$\begin{aligned} \ln(y_{i,t}) - \ln(y_{i,0}) = & -(1 - e^{-\lambda t})\ln(y_{i,0}) + (1 - e^{-\lambda t})\frac{\alpha}{1 - \alpha}\ln(s_i) \quad (15) \\ & -(1 - e^{-\lambda t})\frac{\alpha}{1 - \alpha}\ln(n_i + g + \delta_i) + a_t + u_{i,t} \end{aligned}$$

where $y_{i,0}$ ($y_{i,t}$) is the initial (final) level of real income per worker, λ is the rate of convergence, s is the investment rate, n is labor growth rate, δ is the depreciation rate, and g is the TFP growth rate. The α is the output elasticity of capital in (11). With $\lambda > 0$, the Solow model predicts that the coefficients for $y_{i,0}$ and $(n + g + \delta)$ should be negative, while that for s should be positive.

Our data is comprised of six 10-year panels running from 1950 to 2010. We use personal income minus transfer payments deflated by the regional CPI as real income. We then divide real income by employment to get y . The investment and depreciation rates are taken from our combined series and are averaged across each decade. The labor growth rate n is measured as employment growth during the decade.

We estimate (15) using panel static (FE, RE or FD) and panel SGMM and GMM estimators. The GMM estimators control for the endogeneity of initial real income and s and δ using lagged levels and first-differences. Neither the dynamic mean group (MG, PMG, DFE) nor the SFA estimators are appropriate given the ten-year period for each panel. There is however potential bias arising from spatial externalities in the growth processes of U.S. states (Rey and Montouri, 1999), which we address below.

5.2 Spatial Durbin Model

We use a panel spatial Durbin model (SDM) to control for spatial externalities across states. The spatial Durbin model controls for omitted variables that are spatially correlated to the included regressors, which is likely for regional growth

(LeSage and Fischer, 2008). In our SDM, spatial spillovers can arise from lagged growth, initial income, and investment (and depreciation) rates. We use inverse distance squared as spatial weights and confirm that the SDM is the preferred spatial model following the testing procedure of Elhorst (2010).

6 Results and Discussion

Our focus here is **not** to find the most preferable estimator but rather to assess the plausibility of our data by testing if the estimated values match their theoretical counterparts. As a result, we present our estimates in graphical form in the text and in table format in the appendix. For the unrestricted version, we present confidence bounds of the two coefficients for each estimator along with their predicted value. For the restricted version, we show the confidence intervals of the estimated capital elasticity along with the one-third prediction. We also present a model averaging estimate that combines the previous results into a distribution. Since we aggregate over very different models (some likelihood based and some GMM), we use a simple equal weights averaging since there is no common goodness of fit.

6.1 Production function

Figure 5 presents the 90% confidence bound of each estimator of the unrestricted production function (see Table A1 for details). The coefficient for the labor elasticity β is plotted on the vertical axis and the capital elasticity α is on the horizontal axis. Estimators yielding high uncertainty are in the background, while those generating smaller confidence bounds are in the foreground in darker shade. The results of the unrestricted estimation are for the most part reasonable, although they show mild evidence for increasing returns to scale. The capital elasticity α is generally estimated to be between 0.2 and 0.4, while labor elasticity β mostly range from 0.7 to 1.0. The one notable exception is the MG estimator which is far outside the economically reasonable range, but also estimated with a huge uncertainty which basically renders it uninformative. The solid black line shows the location of estimates with constant returns to scale ($\beta + \alpha = 1$). The constant returns to scale line passes through the confidence bounds for most estimators. The exceptions are the highly precise estimates of PMG, SFA and SGMM1.⁵

Figure 6 shows the 67% and 90% confidence bounds of the capital elasticity α estimates of the restricted production function (see Table A2 for details). The solid horizontal line reflects the null hypothesis of $\alpha = 1/3$. Most of the 90% confidence bounds intersect the null hypothesis, indicating evidence in favor of $\alpha = 1/3$. Those confidence bounds that do not intersect, are usually fairly narrow and located very close to the Null in absolute terms. The poor performance

⁵Contrarily, a Wald test for equality of $\beta + \alpha = 1$ typically rejects. However, the visual results show that this rejection is driven by economically-irrelevant deviations. See Table A1 for details

Note: Each shaded areas represents a 90% confidence bounds of the estimators used in Table A1. Estimators yielding high uncertainty are in the background, while estimators generating smaller confidence bounds are presented in the foreground in darker shades. The bold back line shows the location of estimates with constant returns to scale. The gray area in the back that extends beyond the graph area is the MG estimator.

Figure 5: Confidence bounds of estimated labor and capital elasticity of the unrestricted production function

of GMM can be attributed to weak instruments under highly-persistent capital per worker data. The grew shaded area on the right denotes the model averaging estimator MAvg (including all estimators) and MAvg2 (all but difference GMM, GMM1, GMMc). The 90% confidence bounds for MAvg and MAvg2 contain the predicted one-third value, confirming the plausibility of capital stock data.

6.2 Solow growth model

Figure 7 presents the 90% confidence bound of each estimator of the unrestricted Solow model (see Table A3 for details). The coefficient for $\ln(s)$ is plotted on the vertical axis and the coefficient for $\ln(n+g+\delta)$ is on the horizontal axis. As before, estimators yielding high uncertainty are in the background, while those generating smaller confidence bounds are in the foreground in darker shade. The solid black line shows the location of estimates where the two coefficients sum to zero. The only set of estimates that reject the null are system SGMM

Note: Each entry shows the 67% and 90% confidence bounds of the estimators used in Table A2. The grey shaded area marks the model averaging estimators MAvg (including all individual estimators) and MAvg2 (all but GMM, GMM1 and GMMc). The solid black line reflects the null hypothesis of a capital elasticity of output equal to one-third.

Figure 6: Confidence bounds of estimated capital elasticity of the restricted production function

Note: Each shaded areas represents a 90% confidence bounds of the estimators used in Table A3. Estimators yielding high uncertainty are in the background, while estimators generating smaller confidence bounds are presented in the foreground in darker shades. The bold back line shows the location of estimates where the two coefficients sum to zero.

Figure 7: Confidence bounds of estimated coefficients of the unrestricted Solow Model

shown in dark and FD behind it.

Figure 8 shows the 67% and 90% confidence bounds of the implied capital elasticity α estimates of the restricted Solow model (see Table A4 for details). As before the solid horizontal line reflects the null hypothesis of $\alpha = 1/3$. Most of the estimates of α implied by the Solow growth model roughly match expectations, although they are on average slightly lower. In the Solow model, system GMM generates estimates of α close to zero, while differenced GMM provide estimates closer to one-third. The grey shaded area on the right denotes the model averaging estimator MAvg (including all estimators) and MAvg2 (all but SGMM, SGMM1, SGMMc). The 67% confidence bounds for MAvg and MAvg2 contain the predicted one-third value, but the 90% bounds fall below.

7 Conclusion

This paper presents a new dataset of U.S. state-level capital stock and investment data. We improved the methodology by using output weights in mining

Note: Each entry shows the 67% and 90% confidence bounds of the estimators used in Table A4. The grey shaded area marks the model averaging estimators MAvg (including all individual estimators) and MAvg2 (all but GMM, GMM1 and GMMc). The solid black line reflects the null hypothesis of a capital elasticity of output equal to $1/3$.

Figure 8: Confidence bounds of estimated capital elasticity of the restricted Solow Model

and extended extending the data up to 2015. We intend to make updates of the data available annually when the underlying data is released.

References

- Aigner, D., Lovell, C. K., Schmidt, P., 1977. Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6 (1), 21–37.
- Anderson, T. W., Hsiao, C., 1982. Formulation and Estimation of Dynamic Models using Panel Data. *Journal of Econometrics* 18 (1), 47–82.
- Arellano, M., Bond, S., 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies* 58 (2), 277–297.
- Arellano, M., Bover, O., 1995. Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics* 68 (1), 29–51.
- Battese, G. E., Coelli, T. J., 1995. A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function Panel Data. *Empirical Economics* 20 (2), 325–332.
- BEA, February 1 2018. Table 1.1.10. Percentage Shares of Gross Domestic Product.
URL <https://www.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=2&isuri=1&1921=survey&1903=14>
- Beenstock, M., Felsenstein, D., Zeev, N. B., 2011. Capital Deepening and Regional Inequality: An Empirical Analysis. *The Annals of Regional Science* 47 (3), 599–617.
- Benos, N., Mylonidis, N., Zotou, S., 2016. Estimating Production Functions for the US States: The Role of Public and Human Capital. *Empirical Economics* 52 (2), 1–31.
- Blundell, R., Bond, S., 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* 87 (1), 115–143.
- Blundell, R., Bond, S., 2000. GMM Estimation with Persistent Panel Data: An Application to Production Functions. *Econometric Reviews* 19 (3), 321–340.
- Bode, E., Nunnenkamp, P., 2011. Does Foreign Direct Investment Promote Regional Development in Developed Countries? A Markov Chain Approach for us States. *Review of World Economics* 147 (2), 351–383.
- Brock, G., Ogloblin, C., 2014. Another Look at Technical Efficiency in American States, 1979–2000. *The Annals of Regional Science* 53 (2), 577–590.
- Burnett, J. W., Madariaga, J., 2017. The Convergence of US State-level Energy Intensity. *Energy Economics* 62 (C), 357–370.
- Cardarelli, M. R., Lusinyan, L., 2015. US Total Factor Productivity Slowdown: Evidence from the US States. IMF Working Paper 15/116.

- DiCecio, R., Gascon, C. S., 2010. Income Convergence in the United states: A Tale of Migration and Urbanization. *The Annals of Regional Science* 45 (2), 365–377.
- Elhorst, J. P., 2010. Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis* 5 (1), 9–28.
- Feenstra, R. C., Inklaar, R., Timmer, M. P., 2015. The Next Generation of the Penn World Table. *American Economic Review* 105 (10), 3150–3182.
- Garofalo, G. A., Yamarik, S., 2002. Convergence: Evidence from a New State-by-State Capital Stock Series. *Review of Economics and Statistics* 84 (2), 316–323.
- Griliches, Z., Hausman, J. A., 1986. Errors in Variables in Panel Data. *Journal of Econometrics* 31 (1), 93–118.
- Heckelman, J. C., 2013. Income Convergence among US States: Cross-sectional and Time Series Evidence. *Canadian Journal of Economics* 46 (3), 1085–1109.
- Jacks, D. S., O’Rourke, K. H., Williamson, J. G., 2011. Commodity Price Volatility and World Market Integration since 1700. *Review of Economics and Statistics* 93 (3), 800–813.
- Kalyvitis, S., Vella, E., 2015. Productivity Effects of Public States Maintenance: Evidence from US States. *Economic Inquiry* 53 (1), 72–90.
- LeSage, J. P., Fischer, M. M., 2008. Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spatial Economic Analysis* 3 (3), 275–304.
- Liu, D.-C., 2013. The Evolution of Excess Job Reallocation in the US. *Journal of Macroeconomics* 36 (1), 188–206.
- Magrini, S., 2004. Regional (di) Convergence. *Handbook of Regional and Urban Economics* 4, 2741–2796.
- Mankiw, N. G., Romer, D., Weil, D. N., 1992. A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics* 107 (2), 407–437.
- Metcalf, G. E., 2008. An Empirical Analysis of Eenergy Intensity and its Determinants at the State Level. *The Energy Journal* (3), 1–26.
- Moral-Benito, E., 2013. Likelihood-based Estimation of Dynamic Panels with Predetermined Regressors. *Journal of Business and Economic Statistics* 31 (4), 451–472.
- Mullen, J. K., Williams, M., 2005. Foreign Direct Investment and Regional Economic Performance. *Kyklos* 58 (2), 265–282.

- Nadiri, M. I., Prucha, I. R., 1996. Estimation of the Depreciation Rate of Physical and R&D Capital in the US total Manufacturing Sector. *Economic Inquiry* 34 (1), 43–56.
- Okui, R., 2009. The Optimal Choice of Moments in Dynamic Panel Data Models. *Journal of Econometrics* 151 (1), 1–16.
- Peri, G., 2012. The Effects of Immigration on Productivity: Evidence from US States. *Review of Economics and Statistics* 94 (1), 348–358.
- Pesaran, M. H., Shin, Y., Smith, R. J., 2001. Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics* 16 (3), 289–326.
- Pesaran, M. H., Shin, Y., Smith, R. P., 1999. Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association* 94 (446), 621–634.
- Pesaran, M. H., Smith, R. P., 1995. Estimating Long-run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics* 68 (1), 79–113.
- Reed, W. R., 2008. The Robust Relationship between Taxes and US State Income Growth. *National Tax Journal* (1), 57–80.
- Reed, W. R., 2009. The Determinants of US State Economic Growth: A Less Extreme Bounds Analysis. *Economic Inquiry* 47 (4), 685–700.
- Rey, S. J., Montouri, B. D., 1999. US Regional Income Growth: A Spatial Econometric Perspective. *Regional Studies* 33 (2), 143–156.
- Roodman, D., 2006. How to Do xtabond2: An Introduction to Difference and System GMM in Stata.
- Roodman, D., 2009. A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics* 71 (1), 135–158.
- Sharma, S. C., Sylwester, K., Margono, H., 2007. Decomposition of Total Factor Productivity Growth in US States. *The Quarterly Review of Economics and Finance* 47 (2), 215–241.
- Solow, R. M., 1956. A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics* 70 (1), 65–94.
- Turner, C., Tamura, R., Mulholland, S. E., 2013. How Important are Human Capital, Physical Capital and Total Factor Productivity for Determining State Economic Growth in the United States, 1840–2000? *Journal of Economic Growth* 18 (4), 319–371.
- Yamarik, S., 2008. Estimating Returns to Schooling from State-Level Data: A Macro-Mincerian Approach. *The BE Journal of Macroeconomics* 8 (1), 1–16.

- Yamarik, S., 2011. Human Capital and State-level Economic Growth: What is the Contribution of Schooling? *The Annals of Regional Science* 47 (1), 195–211.
- Yamarik, S., 2013. State-Level Capital and Investment: Updates and Implications. *Contemporary Economic Policy* 31 (1), 62–72.

Appendix

Specification tests

Production function estimation

FE, RE and SFA For the unconstrained model, the Hausman test strongly rejects the random effects specification. This makes the interpretation of the SFA results slightly problematic since it is essentially a permutation of a random-effects model. However, we find slow but significant convergence, indicating that SFA should be used.

RE is no longer rejected for the constrained model. The results of SFA are clearly within the confidence bounds of both FE and RE with extremely narrow confidence bounds for the SFA analysis itself. Following a logic comparable to the Hausman test, it thus seems that SFA is the clearly best choice out of those three at least for the constrained model.

(S)GMM While the coefficient estimates obtained via SGMM and GMM are partly economically reasonable, the specification tests indicate a range of problems. The Arellano-Bond test for autocorrelation does not indicate first order autocorrelation. Assuming uncorrelated errors in the original model – as required for the validity of second lags of levels as instruments for the first lags of first differences – there should be first order autocorrelation by construction. Thus, although we do not find a significant AR(2) term in the difference equation, there seems to be a problematic error structure rendering a GMM type instrumentation questionable.

In addition, we find that the results are extremely sensitive to the choice of instruments. Using the full set of instruments, second lags only, or collapsing the instrument set yields vastly different results, especially for simple GMM estimators that do not include a level equation.

MG, PMG, DFE Because constraining coefficients to be identical across states improves efficiency of the estimation, if the assumption is met, but creates inconsistency, if it is not, we can use a standard Hausman test to distinguish between MG, PMG, and DFE estimators. All our tests fail to reject the more constrained model (i.e. PMG compared to MG, and DFE compared to PMG). This holds true for both the unrestricted specification of the production function, and the specification where we impose constant returns to scale.

Contrary to most other specifications, constant elasticity is not rejected for any of our cointegration estimators. However, in the case of the most general MG specification this is admittedly partly driven by extremely wide confidence bounds. However, out of those three only in the case of our preferred DFE specification, we do not reject a capital elasticity of production that is equal to one-third.

Solow model estimation

FE, RE For both the unconstrained and the constrained Solow model, the RE specification is strongly rejected, indicating that the (economically also more reasonable) fixed effects specification is appropriate.

(S)GMM Contrary to the production function estimation, the specification tests for the GMM models are largely acceptable. We find autocorrelation of order 1, but not of higher order as intended. However, the overidentifying restrictions tests are inconclusive for the most part, with the (non robust) Sargan test rejecting and the Hansen test (which might have power issues) failing to reject.

SDM The Moran's I-statistics for the log of real income per worker and its growth rate are positive and strongly rejects the null of no spatial autocorrelation at any reasonable level. As with the static panel estimators, the SDM-RE specification is strongly rejected, indicating that SDM-FE is preferred.

Table A1: Estimation results of unrestricted production function

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	FE	RE	FD	SGMM	SGMM1	SGMMc	GMM	GMM1	GMMc	MG	PMG	DFE	SFA
α	0.277*** (0.045)	0.302*** (0.053)	0.212*** (0.030)	0.334*** (0.069)	0.289*** (0.083)	0.326*** (0.123)	0.297*** (0.079)	0.122 (0.120)	0.151 (0.126)	-0.166 (1.249)	0.493*** (0.018)	0.387*** (0.048)	0.310*** (0.016)
β	0.964*** (0.089)	0.799*** (0.052)	0.961*** (0.080)	0.768*** (0.081)	0.789*** (0.085)	0.860*** (0.096)	0.805*** (0.205)	1.182*** (0.182)	1.184*** (0.201)	1.171 (1.621)	0.363*** (0.053)	0.559*** (0.131)	0.775*** (0.021)
No of obs	864	864	816	864	864	864	816	816	816	768	768	768	864
States	48	48	48	48	48	48	48	48	48	48	48	48	48
Time Period	1998-15	1998-15	1999-15	1998-15	1998-15	1998-15	1999-15	1999-15	1999-15	2000-15	2000-15	2000-15	1998-15
State Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
$\chi^2 (\alpha + \beta = 1)$	21.96	22.87	8.457	8.426	3.774	5.950	0.342	3.052	3.225	0.229	7.889	0.000	40.550
p-value	0.002	0.000	0.045	0.004	0.052	0.015	0.559	0.081	0.073	0.996	0.005	0.632	0.000
Hausman test		24.05											
p-value		0.000									0.15	0.000	
No. of instruments				318	78	80	289	49	49		0.924	0.999	
overid test				34.60	34.66	30.20	31.11	28.84	27.95				
p-value				1.000	0.994	1.000	1.000	0.526	0.573				
AR(2) test				-0.41	-0.41	-0.43	-0.39	-0.15	-0.20				
p-value				0.682	0.721	0.666	0.693	0.881	0.843				

Note: Standard errors in parentheses where *** p<0.01, ** p<0.05, * p<0.1. FE = fixed effects, RE = random effects, FD = first difference, (S)GMM = (system) GMM with all lags as instruments, (S)GMM1 = (system) GMM with 1 lag as instrument, (S)GMMc = (system) GMM with collapsed instruments, PMG = pooled mean group, DFE = dynamic fixed effects, MG = mean group, SFA = stochastic frontier analysis.

Table A2: Estimates of restricted production function

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
α	FE	RE	FD	SGMM	SGMM1	SGMMc	GMM	GMM1	GMMc	MG	PMG	DFE	SFA
	0.315*** (0.082)	0.313*** (0.647)	0.224*** (0.373)	0.450*** (0.140)	0.373*** (0.149)	-0.308 (0.237)	0.268*** (0.760)	-0.057 (0.219)	-0.020 (0.191)	0.525*** (0.112)	0.426*** (0.167)	0.552*** (0.428)	0.313 (0.160)
No of obs	864	864	816	864	864	864	816	816	816	768	768	768	864
States	48	48	48	48	48	48	48	48	48	48	48	48	48
Time Period	1998-15	1998-15	1990-15	1998-15	1998-15	1998-15	1998-15	1998-15	1998-15	2000-15	2000-15	2000-15	2000-15
State Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
χ^2 ($\alpha = 1/3$)	0.097	0.070	16.050	0.337	0.085	7.250	0.780	3.370	3.354	3.958	136.1	5.027	1.144
p -value	0.827	0.793	0.007	0.562	0.771	0.007	0.377	0.066	0.067	0.467	0.000	0.250	0.285
Hausman test		0.250									0.040	0.000	
p -value		0.618									0.837	0.957	
No. of instruments				168	48	49	153	33	33				
overid test				39.80	38.85	35.80	37.21	23.84	21.71				
p -value				1.000	0.104	0.215	1.000	0.068	0.116				
$AR(2)$ test				-0.40	-0.38	1.09	-0.26	0.66	0.59				
p -value				0.689	0.703	0.215	0.797	0.312	0.555				

Note: Standard errors in parentheses where *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. FE = fixed effects, RE = random effects, FD = first difference, (S)GMM = (system) GMM with all lags as instruments, (S)GMM1 = (system) GMM with 1 lag as instrument, (S)GMMc = (system) GMM with collapsed instruments, PMG = pooled mean group, DFE = dynamic fixed effects, MG = mean group, SFA = stochastic frontier analysis.

Table A3: Estimates of unrestricted Solow model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	(FE)	(RE)	(FD)	(SGMM)	(SGMM1)	(SGMMc)	(GMM)	(GMM1)	(GMMc)	(SDM-FE)	(SDM-RE)
y_0	-0.040*** (0.003)	-0.029*** (0.004)	-0.078*** (0.005)	-0.029*** (0.006)	-0.027*** (0.009)	-0.033*** (0.009)	-0.049*** (0.008)	-0.072*** (0.018)	-0.050* (0.026)	-0.040*** (0.004)	-0.028*** (0.005)
$\ln(s)$	0.010*** (0.002)	0.003* (0.002)	0.008*** (0.002)	0.000 (0.002)	0.000 (0.004)	0.001 (0.005)	0.017** (0.008)	0.019* (0.010)	0.015** (0.007)	0.011*** (0.002)	0.006** (0.002)
$\ln(\Gamma)$	-0.012** (0.006)	0.000 (0.004)	-0.004 (0.006)	0.018*** (0.005)	0.019*** (0.007)	0.003 (0.012)	-0.014 (0.016)	0.011 (0.010)	-0.024 (0.015)	-0.009 (0.006)	-0.002 (0.006)
$W y_0$										0.032** (0.012)	0.012 (0.015)
$W \ln(s)$										-0.010 (0.013)	-0.011 (0.009)
$W \ln(\Gamma)$										-0.016 (0.019)	-0.022 (0.015)
ρ										0.571*** (0.070)	0.559*** (0.075)
σ_e^2										0.000*** (0.000)	0.000*** (0.000)
$\ln t_\theta$										-0.351 (0.293)	-0.351 (0.293)
No of obs	288	288	240	288	288	288	240	240	240	288	288
States	48	48		48	48	48	48	48	48	48	48
$\chi^2(s + \Gamma = 0)$	0.248	0.880	0.610	12.54	7.715	0.161	0.078	6.675	0.519	0.089	2.794
p-value	0.619	0.348	0.435	0.000	0.005	0.688	0.781	0.010	0.471	0.766	0.095
Hausman test		50.34								64.15	
p-value		0.000								0.000	
No. of instruments				48	26	26	35	13	13		
overid test				41.92	30.70	30.20	41.28	14.51	12.26		
p-value				0.345	0.022	0.059	0.039	0.013	0.031		
AR(2) test				1.36	1.46	0.44	0.35	2.17	-0.17		
p-value				0.175	0.146	0.663	0.728	0.030	0.713		

Note: Standard errors in parentheses where *** p<0.01, ** p<0.05, * p<0.1. $\Gamma = n + g + \delta$. W denotes the spatial weight matrix where $w_{i,j} = 1/\text{distance}(i,j)$. FE = fixed effects, RE = random effects, FD = first difference, (S)GMM = (system) GMM with all lags as instruments, (S)GMM1 = (system) GMM with 1 lag as instrument, (S)GMMc = (system) GMM with collapsed instruments, PMG = pooled mean group, DFE = dynamic fixed effects, MG = mean group, SDM = spatial dynamic model.

Table A4: Estimates of restricted Solow model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	(FE)	(RE)	(FD)	(SGMM)	(SGMM1)	(SGMMc)	(GMM)	(GMM1)	(GMMc)	(SDM-FE)	(SDM-RE)
y_0	-0.040 *** (0.003)	-0.029 *** (0.004)	-0.078 *** (0.005)	-0.021 *** (0.005)	-0.026 *** (0.008)	-0.035 *** (0.009)	-0.048 *** (0.008)	-0.075 *** (0.017)	-0.056 *** (0.024)	-0.040 *** (0.004)	-0.029 *** (0.005)
$\ln(s - \Gamma)$	0.010 *** (0.002)	0.003* (0.002)	0.007** (0.003)	-0.001 (0.001)	0.001 (0.004)	0.001 (0.005)	0.018 (0.008)	0.012** (0.006)	0.015** (0.008)	0.010*** (0.002)	0.005** (0.002)
ρ										2.263*** (0.039)	0.561*** (0.073)
$\mathcal{W}y_0$										0.096*** (0.015)	0.006 (0.014)
$\mathcal{W}\ln(s - \Gamma)$										-0.035*** (0.013)	-0.008 (0.009)
Implied α	0.199 *** (0.046)	0.088* (0.048)	0.085*** (0.033)	-0.053 (0.085)	0.024 (0.162)	0.032 (0.132)	0.268*** (0.085)	0.142* (0.080)	0.210* (0.110)	0.205*** (0.046)	0.158*** (0.062)
No of obs	288	288	240	288	288	288	240	240	240	288	288
States	48	48		48	48	48	48	48	48	48	48
$\chi^2(\alpha = 1/3)$	4.869	19.640	41.660	17.400	2.885	4.121	0.143	3.928	0.664	4.173	5.279
p-value	0.027	0.000	0.000	0.000	0.089	0.042	0.706	0.048	0.415	0.041	0.022
Hausman test		50.81								68.15	
p-value		0.000								0.000	
No. of instruments				48	23	29	35	13	13		
overid test				42.05	27.03	32.38	41.44	13.22	13.76		
p-value				0.382	0.028	0.054	0.049	0.040	0.032		
AR(2) test				0.34	0.32	0.21	0.17	0.41	0.12		
p-value				0.736	0.750	0.837	0.867	0.684	0.906		

Note: Standard errors in parentheses where *** p<0.01, ** p<0.05, * p<0.1. $\Gamma = n + g + \delta$. $\mathcal{W}$ denotes the spatial lag operator, where $w_{i,j} = 1/\text{distance}(i,j)$. FE = fixed effects, RE = random effects, FD = first difference, (S)GMM = (system) GMM with all lags as instruments, (S)GMM1 = (system) GMM with 1 lag as instrument, (S)GMMc = (system) GMM with collapsed instruments, PMG = pooled mean group, DFE = dynamic fixed effects, MG = mean group, SDM = spatial dynamic model.