Explaining economic growth with effective area

Octavio Escobar
ESG Management School
25, rue Saint Ambroise
75011 Paris, France
oescobar@esg.fr

Régis Chenavaz
Euromed Management
Domaine de Luminy, Rue Antoine Bourdelle
13288 Marseille, France
regis.chenavaz@euromed-management.com

Abstract

Total area is the usual measure for land factor. As land is unequally useful, total area is a flawed measure. To take into account for land features, we develop an alternative measure called effective area. Effective area is based on spatial population distribution which captures both natural conditions and human activity. We calculate the effective area for 49 states of the United States over 1970-2010 to illustrate the differences between total and effective area. Since effective area is dynamic over time, it is richer from a theoretical perspective than total area. We extend the Solow-Swan model to illustrate the influence of effective area on economic growth. We estimate the model using panel data for the United States for the period 1970-2010. Effective area explains better economic growth than total area. Moreover, it also reduces bias when estimating total factor productivity.
Keywords: Effective area, land, economic growth, United States.
Code JEL: C82, O47, O51, R14.

1 Introduction

Economic activity is linked to capital, labour, and land. Land is an important factor as it provides a place for firm to produce - agriculture, industry, service - and for consumer to live. The measure of land factor is therefore an issue for empirical economics such as international of geographic economics. Empirical literature simply take the total area (or total surface area) as a measure of land without questioning that measure. But the usual measure of total area for land factor is flawed since there is much variety in land quality. In effect some land are more favourable than other. For example, a forest, a desert, access to the sea, landscape, weather, wildlife, oil, fertile soil, and rainfall impact the value of land. Just taking the total area as a measure of land does not take into account for the qualitative aspect of land. How is worth a large land that has extreme climatic and geographic conditions and where no people can live in? In the United States for example, it is difficult to compare the quality of land of two different states such as Arizona and Florida only using their total area. Economic analysis would gain from a better measure for land factor. We propose an alternative measure to land called *effective area*. We define effective area as an area directly useful for production and living at the date under consideration.

We suppose that the quality of the land depends on both natural conditions (or first-nature) and on human activity (or second nature). Natural conditions such as resources, landforms, and geographical location are important factors for people to consider settling in a region. Indeed, people tend to settle in places that have the resources they need to survive and thrive. In addition to natural conditions, human activity influences the attraction of a region through the agglomeration and dispersion forces (Krugman, 1991; Helpman, 1998). In an agglomeration, people benefit from more employment opportunities and lower product prices, while firms benefit from larger employment area and better market access, reflecting an agglomeration force. But not all people and firms agglomerate into a single region since the total area of a region is a limited resource. Ironically, as the population grows, the natural features that may have attracted people to the region are lost or diminished. Moreover, land price increases with population and economic activity, reflecting a dispersion force (Helpman, 1998).

Following Chenavaz and Escobar (2012), this article introduces the effective area estimator. The effective area estimator is based on the Gini coefficient. The effectiveness of an area is linked to its spatial population distribution. If any part of an area is effective, the population is uniformly distributed. Similarly, the lack of population shows and implies that this surface is not effective. We assume that the spatial population distribution captures both natural conditions and human activity. Hence, we can omit detailed data on the underlying characteristics of land such as climate, altitude, coastline, or ground quality. To estimate the effective area of any type of territory, we only need data on population and area for the territory's subdivisions.

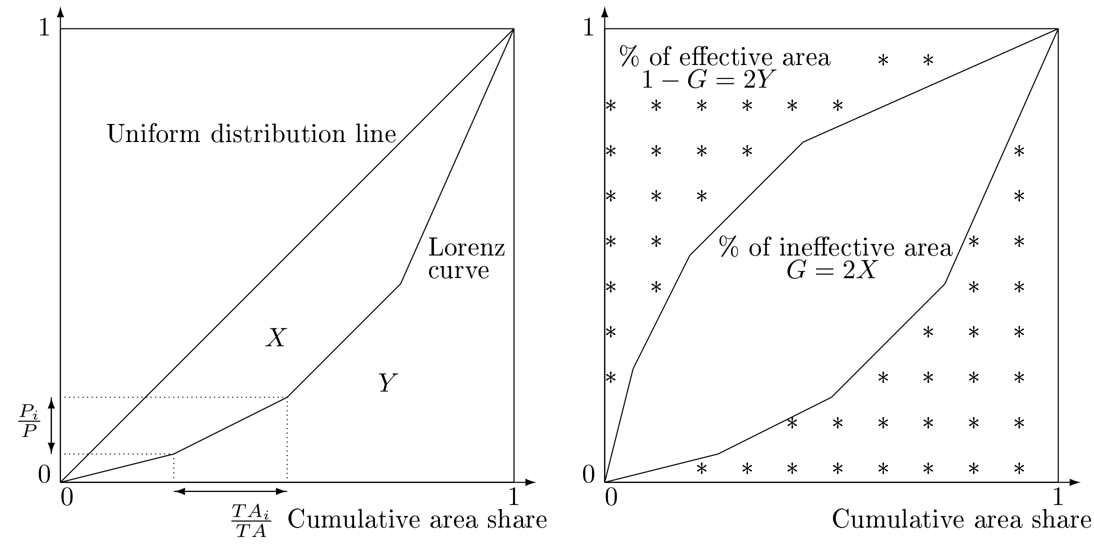
We then apply the effective area estimator to the United States for the period 1970-2010. Results show that there are important differences among population dense states and weather unfriendly states in terms of their proportion of effective area (or effective area as percent of total area). Since variations in natural conditions or in human activities influence the proportion of effective area, the effective area changes over time. Results illustrate that, between 1970 and 2010, some states experienced an increase in their effective area, while others experienced a decrease.

We extend the neoclassical Solow-Swan model to integrate the dynamics of effective area. The model shows that, contrary to total area which is time invariant, shifts in effective area may influence economic growth. In addition, from a theoretical point of view, including effective area into the Solow-Swan model also corrects the measure of total factor productivity. We empirically estimate the model using data for 49 states of the United States for the period 1970-2010. Results show that increasing the effective area improves economic growth. Results also show difference the corrected TFP growth is bigger than the classical TFP growth for most of the states.

2 Effective area estimator

This section follows Chenavaz and Escobar (2012). The Lorenz curve in Figure 1 a represents the spatial cumulative distribution of population. This Lorenz curve connects the cumulative share of population to the cumulative share of area. It is similar to the classical Lorenz (1905) curve which links the cumulative

Figure 1. Spatial cumulative distribution of population and % of effective area



share of wealth to the cumulative share of population.

Let a country be composed of n regions. The regions are indexed by $i \in [1, n]$ ordered by density (or population density). Let the Total Area of the country be TA, the Total Area of region i be TA_i, the Population of the country be P, and the Population of region i be P_i.

The surface between the uniform distribution line - the bisectrix - and the Lorenz curve is $X \in [0, 1/2[$. The surface below the Lorenz curve is $Y \equiv 1/2 - X$:

$$Y = \sum_{i=1}^{n} \frac{TA_i}{TA} \sum_{j=1}^{i} \frac{P_{j-1} + P_j}{2P}, \text{ with } P_0 = 0. \tag{1}$$

The population distribution among the regions is measured by $G \equiv 2X$, which is similar to the classical Gini coefficient. More precisely G is the proportion of ineffective area (Figure 1 b). On the contrary the proportion of effective area is $1 - G \equiv 2Y$:

$$1 - G = \sum_{i=1}^{n} \frac{TA_i}{TA} \sum_{j=1}^{i} \frac{P_{j-1} + P_j}{P}, \text{ with } P_0 = 0. \tag{2}$$

If the Lorenz curve is far from the uniform distribution line, then only a small part of the country's area is effective and the population concentrate in this part; X tends to $1/2$, Y tends to 0, and the proportion of effective area $1 - G$ tends to 0. Alternatively, if the Lorenz curve is on the uniform distribution line, then all the country's area is effective and the population is uniformly distributed; X equals 0, Y equals $1/2$, and the proportion of effective area $1 - G$ equals 1.

The Effective Area is $EA \equiv TA(1 - G)$:

$$EA = \sum_{i=1}^{n} TA_i \sum_{j=1}^{i} \frac{P_{j-1} + P_j}{P}, \text{ with } P_0 = 0. \tag{3}$$

According to (3), if all the population lives in one region i ($P = P_i$), the effective area is region i's total area ($EA = TA_i$). Alternatively, if the population is uniformly distributed in all the regions ($\frac{P_i}{TA_i} = \frac{P_{i+1}}{TA_{i+1}}$), the effective area is the country's total area ($EA = TA$). The effective area has the property of being equivalent to the total area of an equally distributed population. Then the effective area measures the equally distributed population. It is similar to Sen's (1976) welfare index, which measures the mean equally distributed income.

3 Effective area estimation for the United States

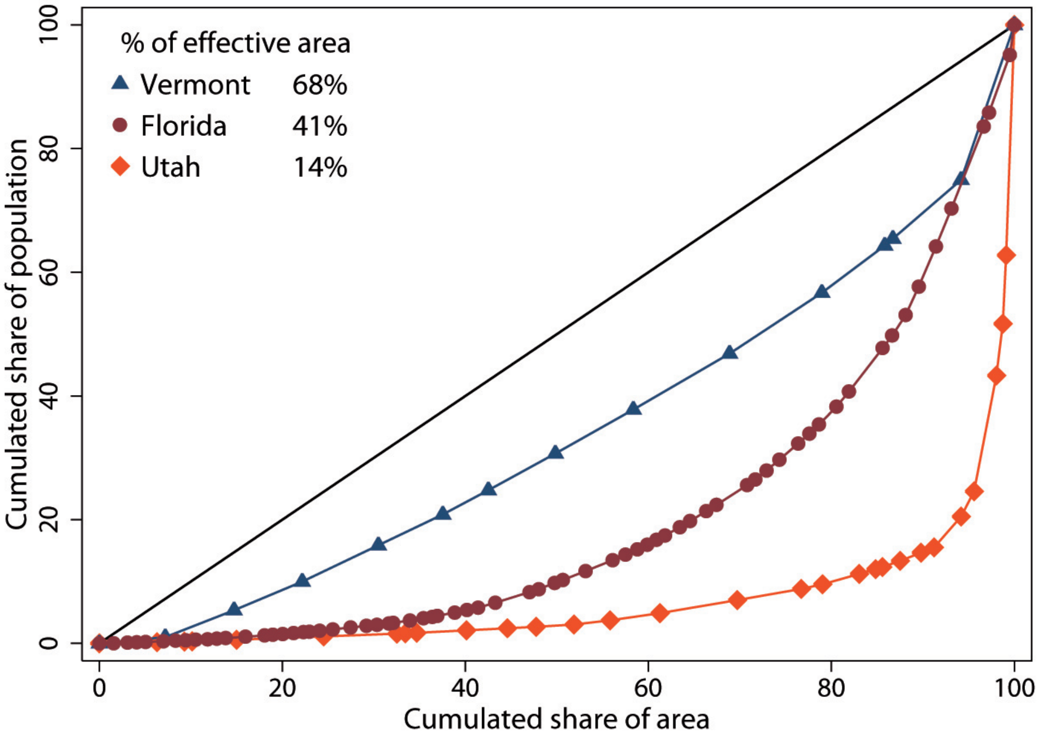
3.1 The data

We estimate the effective area for the United States (US) using yearly panel data for the period 1970-2010. To compute the effective area for each state, we use data on population and area for each state's subdivisions. The US Department of Commerce provides these data at a county-level for each year of the period. Counties include regular counties and county equivalents, such as the parishes of Louisiana, and the independent cities of Maryland, Missouri, Nevada, and Virginia. Population data are from the Bureau of Economic Analysis and counties area data are from the Census Bureau. The number of counties changed between 1970 and 2010 for Alaska, Colorado, Arizona, and New Mexico. For consistency over time issues, we merge the following counties data: Cibola and Valencia in New Mexico, La Paz and Yuma in Arizona, as well as the counties of Adam, Boulder, Broomfield, Jefferson, and Weld in Colorado. Additionally, the Bureau of Economic Analysis combines data for Kalawao County, Hawaii and the small independent cities of Virginia - generally those with fewer than 100,000 residents in 1980 - with those for adjacent counties. Alaska's subdivisions suffer important variations during the period of study. Moreover, the average size of counties in Alaska (70,000 km^2) is considerably bigger than the average size for the other states (2,500 km^2). For these reasons we exclude Alaska from the sample. Finally, we also take out the District of Columbia from the sample because it does not have territorial subdivisions not allowing to compute the effective area.

3.2 Effective area results

According to Figure 2, the relative shapes of the Lorenz curves are coherent with the relative proportions of effective area: 14% for Utah, 41% for Florida, and 68% for Vermont. More precisely, 85% of the population in Utah live in the last total area decile (or the 10% of the total area with highest density). Similarly, in the last total area decile live 41% of the population in Florida and 30% of the population

Figure 2. Lorenz curve and proportion of effective area for selected states



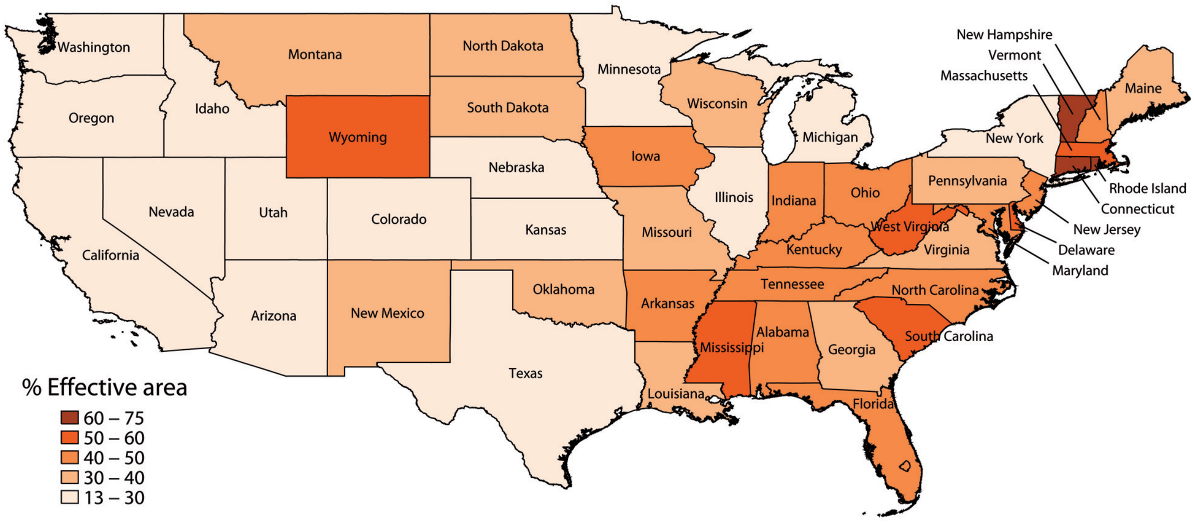
Notes. The Lorenz curve and the proportion effective area are for the year 2010. For each state, the number of segments is the number of counties. The relative size of each segment corresponds to the relative total area and to the relative population of each county.

in Vermont.

The relative proportions of effective area make sense since Utah is one of the most extensive states, one with the lowest population, and a landlocked state with a variety of unfriendly landscapes ranging from arid desert to inaccessibly high elevation areas. Florida is medium size and temperate state, and is one of the most populated states in the US. Its geography is marked by a coastline and much of the state is at the sea level. However, Florida is the most hurricane-prone state. Vermont small and temperate. It is a landlocked state without important agglomerations. Vermont's capital and Vermont's largest city are the least populated state capital and the least populated largest city of a state in the country. Moreover, there is not differences in population size among its cities and its bigger towns. We would obtain similar results if we had taken into account the underlying characteristics of these states such as climate, altitude, terrain variability, or access to the sea.

Since the proportion of effective area depends on geographical characteristics, we may observe similar patterns among neighbouring states. Figure 3 illustrates the proportion of effective area for the different states. We can see different patterns according to the geographical location of the states. While most of eastern states have a proportion of effective area higher than 50%, the proportion of effective area is lower than 30% for most of the western states. We can think that it is because western state have a total area

Figure 3. Proportions of effective area in the United States, 2010



larger than the eastern states. Even within western states, we observe similarities among neighbouring states. For example, among western states, Montana, North Dakota, South Dakota, and Wyoming have a proportion of effective area higher than 30% and all of share a common border. Moreover, these four states have in common an important agricultural industry. New York city, Los Angeles, and Chicago are by far the largest urban agglomerations in the United States; the states hosting these cities have a proportion of effective area lower than 30%. Hence, urbanisation seems to have an influence on the proportion of effective area.

Table 1 reports the estimates of the effective area for each state between 1970-2010. In 2010, Utah, Nevada, and Texas have a proportion of effective area lower than 20%. Rhode Island, Vermont, and Connecticut are the states with the highest proportion of effective area; they have a proportion bigger than 60%. Ranking states by total area and by effective area, more than half of the states have a difference of at least five ranks when comparing the rankings. Although Nevada and Utah total areas are more than twice bigger than those of Kentucky and Ohio; Kentucky and Ohio have more effective area than Nevada and Utah. While in the arid Nevada and in the mountainous and arid Utah population concentrates in some counties - for example, 37% of the Utah's population live in the county of Salt Lake which represents less than 1% Utah's territory - population disperses across the territory on the fertile Kentucky and of the plains of Ohio.

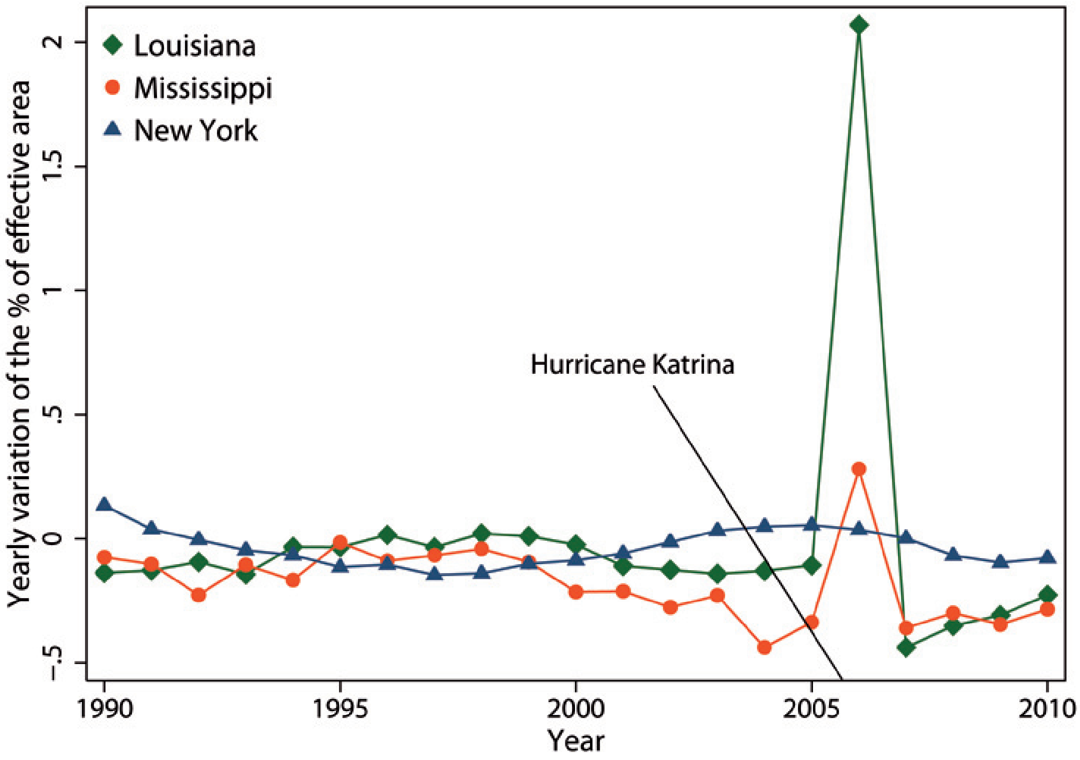
The effective area estimator is based on spatial population distribution, hence a change in population distribution influences its value. In other words, the land factor, measured by the effective area, may vary over time. The quantity of land directly useful shifts with variations in natural conditions and human activities. Table 1 shows the proportion of effective area for the US states for the years 1970, 1980, 1990, and 2010. Delaware, North Dakota, and South Dakota have a variation of 10 percentage

Table 1. Effective area for the United States

State	Counties	Total area	Proportion of Effective Area					Eff. Area
			1970	1980	1990	2000	2010	2010
Alabama	67	131.2	48%	49%	48%	47%	46%	60.1
Arizona	14	294.2	32%	33%	32%	30%	29%	86.6
Arkansas	75	134.8	56%	56%	54%	52%	48%	64.2
California	58	403.5	20%	22%	24%	25%	26%	104.0
Colorado	60	268.4	21%	21%	21%	21%	20%	53.8
Connecticut	8	12.5	60%	61%	62%	61%	62%	7.8
Delaware	3	5.0	49%	52%	53%	56%	60%	3.0
Florida	67	138.9	34%	36%	38%	39%	41%	56.5
Georgia	159	149.0	38%	37%	35%	33%	32%	47.6
Hawaii	4	16.6	24%	26%	28%	31%	33%	5.4
Idaho	44	214.0	32%	30%	28%	26%	24%	50.5
Illinois	102	143.8	25%	26%	25%	24%	24%	34.4
Indiana	92	92.8	45%	47%	47%	46%	45%	41.9
Iowa	99	144.7	56%	54%	52%	50%	47%	67.8
Kansas	105	211.8	33%	32%	29%	27%	25%	52.0
Kentucky	120	102.3	48%	50%	49%	49%	47%	48.4
Louisiana	64	111.9	41%	40%	40%	39%	40%	44.2
Maine	16	79.9	43%	42%	41%	38%	38%	30.0
Maryland	24	25.1	33%	37%	39%	41%	42%	10.5
Massachusetts	14	20.2	49%	52%	53%	53%	53%	10.8
Michigan	83	146.4	25%	27%	28%	29%	29%	42.5
Minnesota	87	206.2	33%	33%	30%	28%	27%	55.2
Mississippi	82	121.5	61%	59%	58%	57%	54%	65.6
Missouri	115	178.0	30%	33%	32%	32%	32%	56.6
Montana	56	377.0	42%	40%	39%	36%	34%	129.0
Nebraska	93	199.0	29%	28%	25%	23%	21%	41.0
Nevada	17	284.3	21%	19%	19%	16%	15%	43.0
New Hampshire	10	23.2	51%	50%	49%	48%	48%	11.1
New Jersey	21	19.0	40%	44%	47%	48%	48%	9.2
New Mexico	32	314.2	38%	36%	35%	34%	32%	99.0
New York	62	122.1	19%	20%	21%	20%	20%	24.1
North Carolina	100	125.9	53%	53%	51%	50%	47%	59.8
North Dakota	53	178.7	55%	51%	46%	42%	38%	68.1
Ohio	88	105.8	37%	39%	40%	41%	41%	43.5
Oklahoma	77	177.7	38%	37%	35%	34%	32%	57.5
Oregon	36	248.6	24%	24%	23%	23%	22%	54.3
Pennsylvania	67	115.9	34%	37%	37%	38%	37%	43.3
Rhode Island	5	2.7	70%	71%	72%	73%	73%	1.9
South Carolina	46	77.9	59%	59%	58%	57%	54%	42.2
South Dakota	66	196.4	45%	43%	40%	37%	34%	66.9
Tennessee	95	106.8	43%	45%	44%	45%	44%	46.5
Texas	254	676.6	22%	21%	19%	18%	17%	113.3
Utah	29	212.8	15%	15%	14%	14%	14%	28.8
Vermont	14	23.9	68%	69%	69%	69%	68%	16.2
Virginia	105	102.3	36%	37%	34%	34%	33%	33.3
Washington	39	172.1	29%	30%	28%	28%	28%	48.4
West Virginia	55	62.3	49%	51%	52%	52%	51%	31.7
Wisconsin	71	140.3	34%	36%	36%	36%	35%	49.5
Wyoming	23	251.5	58%	62%	60%	59%	59%	149.3

Notes. Data source: U.S. Department of Commerce - Bureau of Economic Analysis (counties population) and Census Bureau (counties area). The areas are in thousands of km^2. Counties consist of counties and county equivalents, such as the parishes of Louisiana, the independent cities of Maryland, Missouri, Nevada, and Virginia. Data for Kalawao County, Hawaii and the small independent cities of Virginia - generally those with fewer than 100,000 residents in 1980 - are combined with those for adjacent counties. The number of counties changed between 1970 and 2010 for some states. To keep the same number of counties, we merge for the following counties: Cibola and Valencia in New Mexico, La Paz and Yuma in Arizona, as well as Adam, Boulder, Broomfield, Jefferson, and Weld in Colorado. For consistency and homogeneity we exclude Alaska from the sample. Alaska's subdivisions suffer important variations during the period of study. Moreover, the size of counties in Alaska is considerably bigger than the average. We also take out the District of Columbia from the sample because it does not have territorial subdivisions not allowing to compute the effective area.

Figure 4. Yearly variation of the proportion of effective area for selected states over 1990-2010



Notes. We define yearly variation of the proportion of the effective area as $\triangle_t \%EA \equiv \%EA_t - \%EA_{t-1}$.

points or more in their proportion of effective surface between 1970 and 2010. Only, the effective surface of Delaware increases from 49% in 1970 to 60% in 2010. Delaware's geographical location and business-friendly corporation law attract firms and people. Indeed, more than half of the population growth is on account of outsiders, either from other states or form other countries. Results suggest that these population disperses across Delaware's territory. North Dakota is the state with the highest decrease in effective area from 55% in 1970 to 48% in 2010. This is because of low population growth rates for North Dakota during the period, as well as a transformation of the economy from agriculture industry to oil production, technology and service sectors. Table 1 also illustrates that variations in the proportion of effective area in a decade are weak. Indeed, most of the states have a ten years variation in their proportion of effective area lower than 2 percentage points between 1970 and 2010.

Natural disasters or social conflicts may also impact the proportion of effective area. In august 2005, the hurricane Katrina hit the US, particularly Louisiana and Mississippi. Figure 4 highlights the effects of hurricane Katrina on the proportion of effective surface of Louisiana and Mississippi between 2005 and 2006. This hurricane touched New-Orleans and Blix in Louisiana and other towns in Mississippi. As a result, the population dispersed and the proportion of effective area increased. On the contrary, New York was not threatened, the population did not dispersed, and the proportion of effective area holds.

4 Revisiting economic growth theory with effective area

4.1 Adding land factor to the Solow-Swan model

The classical Solow-Swan model uses the variables production Y, capital K, labour L, and a productivity parameter A. Nordhaus (1992) adds the variable land T. Following Nordhaus (1992) and Romer (2011), the Cobb-Douglas production function with values in $\mathbb{R}$ at time t is

$$Y(t) = A(t)K(t)^{\alpha}L(t)^{\beta}T(t)^{\gamma}, \quad (\alpha, \beta, \gamma) \in \mathbb{R}^{3+}. \tag{4}$$

Note that the productivity parameter A tell us how productively the economy employs the factors K, L, and T. Hence, A represents not just labour productivity, but the total factor productivity (TFP).

Hereafter to improve clarity, a dot between variables denotes multiplication and we omit the argument t.

Taking the logs of both sides of (4) gives

$$lnY = lnA + \alpha lnK + \beta lnL + \gamma lnT.$$

There are two cases for the specification of land T. In the first case, land is the total area TA. In the second case, land is the effective area EA.

$$lnY = lnA + \alpha lnK + \beta lnL + \gamma lnTA, \tag{5a}$$

$$lnY = lnA + \alpha lnK + \beta lnL + \gamma lnEA. \tag{5b}$$

The measure of production in log form using total area in ($5a$) or effective area in ($5b$) seems equivalent because in each case four explicative variables appear. But the implications are distinct since the standard hypothesis (Nordhaus, 1992; Romer, 2011) considers land - measured by total area - fixed over time. Therefore, $\dot{TA} = 0$ and thus $g_{TA} = 0$ where $\dot{TA} \equiv \frac{dTA}{dt}$ is the derivative of TA with respect to time t and $g_{TA} \equiv \frac{\dot{TA}}{TA}$ is the growth rate of TA. The new hypothesis concerns the measure of land by effective area. We take into account the dynamics of population distribution in a country. In effect, population can move from one region to another or population growth rates differ among regions. Thus the effective area EA based on spatial population distribution may change. Therefore, $\dot{EA} \in \mathbb{R}$ and thus $g_{TA} \in \mathbb{R}$.

Differentiating equations ($5a$) and ($5b$) with respect to time t and rewriting the equations in terms of growth rates enlighten the richness of using effective area rather than total area.

$$g_Y = g_A + \alpha g_K + \beta g_L, \tag{6a}$$

$$g_Y = g_A + \alpha g_K + \beta g_L + \gamma g_{EA}. \tag{6b}$$

The measures of production in terms of growth rate using total area in $(6a)$ or effective area in $(6b)$ does not seem equivalent any more. In effect, the measure of production growth using the total area only implies three variables. As the total area is fixed, its dynamic vanishes ($\dot{TA} = 0 \rightarrow g_{TA} = 0$). In contrast the measure of production growth using the effective area still implies four variables. As the population distribution varies, the dynamic of the effective area remains ($\dot{EA} \in \mathbb{R} \rightarrow g_{EA} \in \mathbb{R}$). Thus the specification of land is theoretically richer with the effective area than with the total area.

Since total factor productivity (TFP) ranges from technology to human capital, it cannot be measured directly. Hence, economic literature measure it as the residual (also known as the Solow residual) that accounts for the effects in output not caused by production factors. From equations $(6a)$ and $(6b)$, we can estimate the TFP growth for each specification as:

$$g_A = g_Y - \alpha g_K - \beta g_L, \tag{7a}$$

$$g_{A'} = g_Y - \alpha g_K - \beta g_L - \gamma g_{EA}. \tag{7b}$$

Thus

$$g_{A'} = g_A - \gamma g_{EA} \tag{8}$$

From a theoretical point of view, the modified TFP growth $g_{A'}$ is the classical TFP growth g_A minus the output elasticity to effective area γ times the effective area growth g_{EA}. If γ and g_{EA} are different from zero, classical TFP growth is biased by γg_{EA}. More precisely, if $g_{EA} > 0$ (the population repartition is more uniform), g_A overestimates the TFP growth. If $g_{EA} < 0$ (the population repartition is more concentrated), g_A underestimates the TFP growth.

4.2 Empirical specification and data

Following equations (5a) and (5b), empirical specification is

$$lnY_{it} = \alpha\, lnK_{it} + \beta\, lnL_{it} + \gamma\, lnTA_{it} + \varepsilon_{it}, \tag{9a}$$

$$lnY_{it} = \alpha\, lnK_{it} + \beta\, lnL_{it} + \gamma\, lnEA_{it} + \varepsilon_{it}, \tag{9b}$$

where $\varepsilon_{it} = lnA_{it}$ is the error term for TFP.

The sample includes the 49 states of the US for which we compute the effective area in the previous section. Additionally to effective area, we need data on production, labour, capital and total area. We measure production Y as the real Gross Domestic Product in 2005 dollars. State-level data on capital is not available. Hence, to test robustness, we use two different proxies for capital K. The first proxy is the income for interest, dividends, and rents (thereafter capital income) in 2005 dollars. The second proxy is the capital consumption in 2005 dollars. Labour L data includes only employees, not the self-employed. *Total area* is the area in squared-kilometres, and *Effective area* is the measure obtained in the previous section. Production, labour, and capital data are from the Bureau of Economic Analysis. The data are annual and cover the period 1970-2010.

4.3 Empirical results

We use three different estimation techniques: General least squares random-effects estimator (RE), fixed-effects within estimator (FE), and Arellano and Bond (1991) dynamic panel generalized method of moments (DGMM).

Table 2 reports the estimates of equations (9a) and (9b) using capital income as a proxy of capital. Columns 1 to 2 present the results using equation (9a). Random-effects estimates (column 1) show non significant correlation between total area and GDP. An estimate for total area is not possible when controlling for fixed-effects (column 2) because it is a time invariant variable. Capital and labour variables are significant at 1% level in both specifications. Moreover, there is not a significant difference in the coefficients values for capital and labour variables, the various R^2 values are also similar, as well as the root mean square error (RMSE).

Columns 3 and 4 report the estimates using the effective area (equation 9b). Random-effects estimates (column 3) for capital and labour variables are similar to those found previously. Measuring land as effective area gives positive and significant estimates at 1% level. Increasing the effective area in 10 percentage-points leads to an increase in the GDP of 0.5 percentage-point. The estimates suggest a stronger impact of increasing the effective area on the GDP when controlling for state-level fixed-effects. Since the fixed-effects specification controls for state-specific characteristics such as land, we can interpret this result as the influence of an increase in the proportion of effective area on the GDP. In other words, a 10 percentage-points increase in the proportion of effective area increases the GDP in 4.24 percentage-points. Finally, RMSE values suggest that the predicting power is higher when using the effective area and the fixed-effects specification.

Since we use capital income as a proxy of capital, there are potential endogeneity caused by omitted variables or by measurement error. In addition, potential simultaneity among variables may also create endogeneity issues. For example, labour stimulates output, but to increase output firms raise the demand

Table 2. Effects of total area and effective area on economic growth

Dependent variable: log of GDP	(1)	(2)	(3)	(4)	(5)	(6)
	RE	FE	RE	FE	DGMM	DGMM^c
Capital	0.386**	0.391**	0.383**	0.343**	0.101*	0.099
Income	(0.061)	(0.065)	(0.061)	(0.060)	(0.049)	(0.077)
Labour	0.660**	0.676**	0.663**	0.793**	0.433**	0.453**
	(0.071)	(0.103)	(0.072)	(0.101)	(0.095)	(0.083)
Total Area	-0.014					
	(0.014)					
Effective Area			0.050**	0.424**	0.529**	0.578**
			(0.017)	(0.121)	(0.151)	(0.143)
GDP$_{t-1}$					0.664**	0.707**
					(0.036)	(0.051)
Observations	2009	2009	2009	2009	1911	1911
Instruments					492	194
R^2 within	0.976	0.976	0.977	0.979		
R^2 between	0.994	0.994	0.990	0.880		
R^2 overall	0.992	0.991	0.989	0.886		
RMSE	0.060	0.060	0.060	0.056		
Hansen J p-value					1.000	1.000
AR(1) p-value					0.001	0.001
AR(2) p-value					0.684	0.709

Notes. * significant at 5%, ** significant at 1%. All variables are expressed in log form. RE refers to GLS random-effects estimates. FE is the the fixed-effects (within) regression estimates. DGMM is the Arellano and Bond (1991) dynamic panel GMM estimator. DGMM regressions treat the explanatory variables as endogenous. DGMM uses two lags of endogenous variables as instruments. DGMM^c uses a collapsed instrument matrix to reduce the number of instruments. Robust standard errors are in parentheses. Hansen J-test reports the p-values for the null hypothesis of instrument validity. The p-values reported for AR(1) and AR(2) are the p-values for first and second order autocorrelated disturbances. Each regression includes a constant and time dummies not reported here.

of labour. To handle these endogeneity issues, we proceed as follows. First, we added a one-year lag of GDP variable to reduce omitted variable bias. Indeed, one-year lagged GDP also captures the stock of capital in the economy. Second, we use the Arellano and Bond (1991) DGMM estimator to handle endogeneity, time-invariant specific characteristics, and autocorrelation generated by the inclusion of one-year lagged GDP variable.

We consider all explanatory variables as endogenous variables (excepting time dummies). We instrument endogenous variables with their first and second lags in levels. Column 5 shows that including one-year lag of GDP and using DGMM estimator reduces the coefficient value of both capital income and labour variables, but they are still significant. As expected, one-year lag of GDP is significant and positively correlated to GDP. Estimates for the effective area variable are closer from the fixed-effects estimates than from the random-effects estimates. It suggest that an increase of 10% of effective area leads to an increase of around 5% of GDP.

The row for the Hansen J-test reports the p-values for the null hypothesis of the validity of the over-identifying restrictions. We do not reject the null hypothesis of instruments validity. The values reported for AR(1) and AR(2) are the p-values for first and second order autocorrelated disturbances in the first-differenced equation. As expected, there is high first order autocorrelation, and no evidence for significant second order autocorrelation. These test statistics hint at a proper specification.

The Hansen J-test can be weakened by instrument proliferation. We follow Roodman (2009) and also

Table 3. Effects of total area and effective area on economic growth

Dependent variable: log of GDP	(1)	(2)	(3)	(4)	(5)	(6)
	RE	FE	RE	FE	DGMM	DGMM^c
Capital Consumption	0.067**	0.066*	0.066**	0.060*	0.007	0.002
	(0.025)	(0.026)	(0.025)	(0.023)	(0.005)	(0.005)
Labour	0.964**	0.954**	0.962**	1.039**	0.455**	0.450**
	(0.039)	(0.066)	(0.040)	(0.060)	(0.080)	(0.105)
Total Area	-0.022*					
	(0.011)					
Effective Area			0.046	0.450**	0.479**	0.494**
			(0.024)	(0.120)	(0.120)	(0.158)
GDP_{t-1}					0.701**	0.779**
					(0.038)	(0.050)
Observations	2009	2009	2009	2009	1911	1911
Instruments					492	194
R^2 within	0.976	0.976	0.977	0.979		
R^2 between	0.993	0.993	0.988	0.852		
R^2 overall	0.991	0.991	0.987	0.862		
RMSE	0.061	0.060	0.060	0.056		
Hansen J p-value					1.000	1.000
AR(1) p-value					0.001	0.001
AR(2) p-value					0.719	0.764

Notes. * significant at 5%, ** significant at 1%. All variables are expressed in log form. RE refers to GLS random-effects estimates. FE is the the fixed-effects (within) regression estimates. DGMM is the Arellano and Bond (1991) dynamic panel GMM estimator. DGMM regressions treat the explanatory variables as endogenous. DGMM uses two lags of endogenous variables as instruments. DGMM^c uses a collapsed instrument matrix to reduce the number of instruments. Robust standard errors are in parentheses. Hansen J-test reports the p-values for the null hypothesis of instrument validity. The p-values reported for AR(1) and AR(2) are the p-values for first and second order autocorrelated disturbances. Each regression includes a constant and time dummies not reported here.

estimate the model using a collapsed instrument matrix for the DGMM estimator (column 6). Excepting capital income variable which become non significant at 5% level, coefficient values and significance are very similar to those presented in column 5.

To test robustness, we replace capital income variable by capital consumption variable. Table 3 presents the results. Capital consumption variable is weaker to explain the GDP than capital income variable. Indeed, capital consumption coefficient values are lower for all estimators. Moreover, when using DGMM estimator capital consumption coefficient value is near to zero and its standard errors values suggest that it is not significantly different to zero.

RE and FE estimates show that replacing capital variable increases coefficient values of labour variable for arount 50%. Hence, the importance of controlling for endogeneity. DGMM estimates (columns 5 and 6) for labour, effective area, and one-year lagged GDP are quite similar to those found using in the previous specification. For both Table 2 and Table 3 estimates, the coefficient value of labour and effective area variables are around 0.5 when using DGMM estimator, while the coefficient value of one-year lagged GDP is around 0.7.

Table 4. TFP growth and modified TFP growth for the United States, 1970-2010

State	Classical TFP growth	Modified TFP growth	State	Classical TFP growth	Modified TFP growth
Alabama	15.9%	20.4%	Nebraska	22.2%	35.4%
Arizona	20.2%	17.6%	Nevada	1.3%	8.0%
Arkansas	10.9%	18.2%	New Hampshire	36.3%	37.5%
California	29.5%	18.8%	New Jersey	31.0%	24.4%
Colorado	28.6%	27.9%	New Mexico	7.4%	12.9%
Connecticut	51.4%	52.1%	New York	36.6%	37.3%
Delaware	62.3%	52.9%	North Carolina	21.9%	27.9%
Florida	20.4%	9.1%	North Dakota	24.9%	38.3%
Georgia	23.0%	29.9%	Ohio	18.4%	14.7%
Hawaii	15.0%	2.1%	Oklahoma	18.3%	24.4%
Idaho	7.8%	17.4%	Oregon	23.4%	25.1%
Illinois	23.6%	28.2%	Pennsylvania	22.4%	21.2%
Indiana	23.4%	23.5%	Rhode Island	36.1%	37.7%
Iowa	22.9%	29.1%	South Carolina	20.5%	26.0%
Kansas	15.1%	26.7%	South Dakota	21.3%	32.1%
Kentucky	4.1%	4.8%	Tennessee	25.2%	24.9%
Louisiana	13.1%	16.5%	Texas	30.8%	39.8%
Maine	24.1%	30.7%	Utah	19.3%	18.5%
Maryland	25.3%	16.7%	Vermont	10.2%	10.0%
Massachusetts	36.2%	35.2%	Virginia	30.4%	35.1%
Michigan	12.4%	6.7%	Washington	15.6%	15.2%
Minnesota	16.7%	25.0%	West Virginia	10.1%	10.4%
Mississippi	20.1%	27.2%	Wisconsin	22.8%	20.3%
Missouri	16.0%	14.5%	Wyoming	20.0%	17.9%
Montana	-0.3%	7.1%	*Mean*	*21.7%*	*23.5%*

Notes. We use FE estimates reported in Table 2 columns 2 and 5 to calculate classical TFP and modified TFP growth respectively.

4.4 Effective area and total factor productivity

Tables 1, 2, and 3 show that effective area varies ($g_{EA} \neq 0$) and it is significantly correlated to GDP ($\gamma \neq 0$). Hence, from equations 9a and 9b, we deduce that there are differences between classical TFP growth (g_A) and modified TFP growth ($g_{A'}$).

We use FE estimates reported in Table 2 to calculate, for each state, both classical TFP and modified TFP. Classical TFP is estimated using the estimates of the specification without the effective area variable (column 2); while modified TFP is calculate using the estimates of the specification with the effective area variable (column 4). We then estimate classical TFP growth as well as modified TFP growth between 1970 and 2010.

Table 4 presents the classical TFP and the modified TFP growth for the 49 states in the sample, as well as their mean. In average not including effective area in the model undervalue effective factor productivity growth of 1.8 percentage point. Indeed, for most states (30/49), classical TFP growth is weaker than modified TFP growth. For North Dakota, Nebraska, Hawaii, Kansas, and Florida, differences between classical TFP and modified TFP are bigger than 11 percentage points. For Indiana, Vermont, Tennessee, West Virginia, and Washington, these differences are lower than 0.4 percentage point. Finally, one state, Montana, has negative classical TFP growth, but a positive modified TFP growth.

Delaware and Connecticut are by far the states with the highest classical TFP growth and modified

TFP growth. However, when using modified TFP growth the difference between these two states is lower. New York is the third state in terms of classical TFP growth with a growth of 36.6%. However, when using modified TFP growth, Texas becomes the third state with a modified TFP growth of 39.9%. These results highlights the importance of using effective as a production factor.

5 Conclusion

Since land factor is hard to measure, empirical literature use total area to measure land. The total area however biases the results since land characteristics are unequal within territories of same total area. We propose a measure of land factor called effective area. We suppose that population settle in regions with better land characteristics and that population presence influences the attractiveness of those regions. Land characteristics differences across the sub-divisions of a territory lead to different population distribution within this territory. Hence, we use the spatial distribution of population to estimate the effective area of a territory. The effective area estimator summarises detailed data that are time consuming to find. The effective area is also easy to calculate and uses data usually available in the public domain.

We apply the effective area estimator to the United States. Small and population dense states such as Connecticut, Rhode Island, and Vermont have a proportion of effective area higher than 60%; while big and weather unfriendly states such as Nevada, Texas, and Utah have a proportion of effective area lower than 20%. These results would be similar if we had taken into account the climatic and geographic features of the states.

We show that the proportion of effective area is not static. Effective area shifts with variations in natural conditions and human activities. For example, we illustrate how the hurricane Katrina influenced effective area in both Louisiana and Mississippi, but not in states not threatened by the hurricane.

Since effective area varies over time, land factor measured by the effective area may also varies. We propose then to include effective area in the Solow-Swan model. Hence, shifts in effective area influences economic growth. In addition, including effective area into the Solow-Swan model also corrects the measure of total factor productivity.

We empirically estimate the model using data for 49 states of the United States for the period 1970-2010. Contrary to standard total area measure, effective area is positively significant at 5% level for different specifications and estimators. More precisely, increasing effective area of 10% leads to a GDP growth of around 5%. Finally, we compute the total factor productivity and the effective total factor productivity for each state. Between 1970 and 2010, modified TFP growth is bigger than classical TFP growth for most of the states. Montana experienced a decrease in classical TFP, but its factor productivity increases when using modified TFP.

Using effective area instead of total area not only reduces the bias, but it has important policy implications. Indeed, authorities can adopt business and migration friendly policies to attract firms and people to increase their effective area and then their GDP.

References

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies 58*(2), 277–97.

Chenavaz, R. and O. Escobar (2012). Effective area as a measure of land factor. *Economics Bulletin 32*(3), 1962–1969.

Helpman, E. (1998). The Size of Regions. In D. Pines, E. Sadka, and I. Zilcha (Eds.), *Topics in Public Economics: Theoretical and Applied Analysis*, pp. 33 – 54. Cambridge: Cambridge University Press.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483.

Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Publications of the American Statistical Association 9*(70), pp. 209–219.

Nordhaus, W. D. (1992). Lethal model 2: The limits to growth revisited. *Brookings Papers on Economic Activity 23*(2), 1–60.

Romer, D. (2011). *Advanced Macroeconomics*. McGraw-Hill Series Economics. McGraw-Hill Education.

Roodman, D. (2009, March). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal 9*(1), 86–136.