Transitioning out of Poverty*

David Brasington[†], Mika Kato[‡], and Willi Semmler[§]

November 26, 2008

Abstract

We analyze the role of social environment and human capital formation in persistence of poverty and inequality. We present a Romer (1990) type variety model where the presence of economies of agglomeration in social environment may cause two basins of attraction; whereby we may interpret the lower basin as a *poverty trap* and the upper basin as a *take-off* region. The long-run economic status of households and the formation of social environmental capital and human capital crucially depends on its initial social and human resources in the community. We also consider the size of income transfer to regions and its effect on inequality and welfare. We provide supporting evidence of existing inequality and poverty trap using educational attainment data for the U.S..

Key Words: inequality, growth, social environment, human capital, economies of agglomeration

JEL Codes: C61, O15, R11

1 Introduction

Adam Smith in *The Wealth of Nations* in 1776 stated his views on inequality as follows:

> "The difference between the most dissimilar characters, between a philosopher and a common street porter, for example, seems to arise not

*We would like to thank two anonymous referees for their valuable comments and suggestions. We would also like to thank the participants of the 12th International Conference on Computing in Economics and Finance, Limassol, Cyprus (June, 2006), the Department Seminar at American University (January, 2007), and the Conference on Institutional and Social Dynamics of Growth and Distribution, Luca, Italy (December, 2007) for their discussion on an erlier version of this paper. We also want to thank Uwe Köller for technical assistance.

[†]University of Cincinnati; david.brasington@uc.edu

[‡]Howard University, Washington, DC; mkato@howard.edu

[§]New School for Social Research, New York; SemmlerW@newschool.edu

> so much from nature, as from habit, custom, and education. When they came into the world, and for the first six or eight years of their existence, they were, perhaps, very much alike, and neither their parents nor playfellows could perceive any remarkable difference." (Smith, 1852, p.7)

Smith thus argues that 1) social environment such as habit and custom and 2) education play a larger role in shaping individual's future than inborn talents do.

The role of education and human capital formation in inequality, which is the second point by Smith, has been emphasized in many studies (Mincer, 1958; Katz *et al.*, 1999, and Bowles and Gintis, 2002). Recent research on inequality such as Durlauf (2004, 2006) and Brock and Durlauf (2006), on the other hand, emphasizes the role of social interaction for socioeconomic outcomes, which is similar to the first point by Smith. A Brock-and-Durlauf type model derives an important implication that the social environment surrounding an individual can lead to a take-off of the individual or can lead to substantial immobility, the so-called social lock-in. Their model, however, does not explicitly discuss the role of human capital formation. Therefore, we are interested in providing a model that takes both social environment and human capital formation into account and whereby we can analyze what social setting is likely to lead to a take-off or persistent poverty.

We present a Romer (1990) type dynamic model where a typical household in a community faces a decision about consumption and investments in environment and education. A state of a community at a given point of time is described by environmental capital stock and human capital capital. By 'environment', we mean the general attractiveness of a community, for example, as a place for working, educating children, doing business, and social contacts, which is largely determined by the magnitude of public investments as well as private investments. Public investments usually provide the community with the basic public services such as education, health care, public transportation, safety, sanitation while private investments add variety of jobs and services to the community.

The community's environment can improve when public or/and private investments are made although their function and nature are quite different. Social investments can be planned while private investments are mostly exogenous to the community. The latter investments, however, may depend on the community's current environment. This may be so due to the so-called economies of agglomeration; as the community's environment improves, it can attract more private investments and with more private investments its environment can improve further, although such feedback, of course, may slow down in later stages due to congestion, etc.. The planned investment in environment, therefore, can stimulate private investments indirectly.

Our model, under stronger economies of agglomeration, may exhibit two basins of attraction where the lower basin can be interpreted as a poverty trap and the upper basin as a take-off region. There is also a threshold separating the two basins of attraction. The dynamics are open-ended in the sense that a community with initial environmental and human resources above a certain threshold level tends toward the upper (high-income) steady state while a community with those below the threshold tends toward the lower (low-income) steady state. The presence of such a mechanism has important implications for competition in the market and for poverty and aggregate inequality.[1]

In this context we can discuss public policies aimed at reducing inequality. Inequality matters as it may create a number of undesirable outcomes in the society such as strong tensions between the poor and the rich, social instability, and social unrest. We show that increasing the budget given to an individual household specifically for the purpose of revitalizing the community[2] may increase income inequality that is simply measured by the distance between the upper and the lower steady states although it may reduce inequality in terms of welfare. In this sense, inequality in welfare may be a better policy measure than inequality in income.

The rest of the paper is organized as follows. Section 2 provides a brief survey on existing theories and empirics of human capital formation, the role of the social environment and inequality. Section 3 introduces our dynamic model that can explain a mechanism of lock-in as well as that of take-off. With the help of a numerical study, we explore the effect of economies of agglomeration on the local and global dynamics and derive the policy function. Section 4 presents an empirical study of state-dependent transition – one of the model's implications. We use math proficiency data for school districts in Ohio for the period 1990-2002 and study whether the transition of educational attainment as a measure of human capital is affected by the past state of attainment. Data indicate a state-dependent pattern of transition and the existence of a threshold, above which school districts tend to improve toward better categories and below which school districts remain in substantial immobility. Section 5 concludes the paper.

2 Human capital, social environment and inequality

It is often maintained that education and human capital formation are a fundamental force for economic growth and income increase in regions. As human capital accumulates, incomes eventually will rise, and poor regions are likely to transition out of poverty. This section briefly surveys the existing theories of human-capital-led growth and inequality and relates this to the new literature on the role of social environment. We also provide a review on the empirical evidence on these topics.

[1] For a threshold model on inequality across countries, see Semmler and Ofori (2007).

[2] Success stories may include a number of gentrification projects in some areas in the NYC and Washington, DC that began in early 1990s, which eventually created hundreds of new condominiums, many new upscale restaurants, bars, shops, theaters, museums, galleries, and other attractions. This finally also led to lower crime rates. The number of murder in NYC dropped from 2262 in 1990 to 539 in 2005 and the number of homicides in DC dropped from 399 in 1994 to 169 in 2006.

2.1 Theories of inequality

The oldest type of explanation relates inequality to the distribution of the individual abilities. This, however, doesn't really explain why the highly skewed income distribution emerges from the normally distributed inborn abilities. As the above-cited statement by Adam Smith indicates, it is highly doubtful that inborn abilities play a relevant role in explaining the income inequality. Recently economists and social scientists point that the acquired knowledge and skills, and the available resources and the environment are more critical determinants of disparities of income.

Mincer (1958) is a landmark work that relates investment in human capital in a direct way to income inequality. Mincer formalizes an individual's decision on the life time allocation between training and work. As a result of different individual preferences, some choose the combination of shorter training and a low-income job with a longer life-time at work while others choose the combination of longer training and a high-income job with a shorter life-time at work. Therefore, the resulting income distribution, in his discussion, reflects a matter of individual taste and preference.

From the late 1960s to 80s, human-capital-led growth theory made great stride (Uzawa, 1965; Ben-Porath, 1967; and Lucas, 1988). Income inequality in this type of model arises from initial endowments of the individuals in addition to the individual ability of skill development. Examples of what characterizes individual's given endowments are the characteristics of parents and family members as well as those of the group, the community, the region, and the country to which an individual belongs. Becker and Tomes (1979), for example, put great emphasis on the role of family characteristics in human capital formation. Educational choice to improve his or her own abilities in their model is considered as a family's problem, especially a parent's problem, rather than an individual problem. Income inequality, therefore, can emerge due to various degrees of parental altruism toward children and the parents' income stream. Galor and Zeira (1993) and Matsuyama (2000, 2007) emphasizes the role of credit market imperfections in persistent inequality when individuals may borrow to invest in human capital. As borrowing decision in imperfect credit markets depends on the initial wealth inherited, wealthier individuals tend to invest in human capital and thus leave more bequests to their descendants, and vice versa. As a result, wealth inequality may amplify and persist generation after generation. In this context then there will arise a strong likelihood of the inheritance of inequality (see Bowles and Gintis, 2002).

Yet, as above-mentioned, more recent research such as Durlauf (2004, 2006) and Brock and Durlauf (2006) puts forward a social-interaction theory of inequality. It broadly emphasizes the social environment based on the conjecture that the composition of groups to which a person belongs plays an important role for socioeconomic outcomes. This is, as Brock and Durlauf argue, because individual preference, beliefs and opportunities are strongly shaped and impacted by one's membership in a particular group (e.g., in a neighborhood, a school, a university, and a workplace). When positive interaction effects occur in a certain group or in a social environment, this can give rise to better opportunities to the group members and create common or similar outcomes for the group members, but this may also cause a greater cross-sectional inequality and less social mobility, i.e., a considerable lock-in, unless a take-off can take place.

2.2 Empirical literature

The formation of human capital is usually measured from the input side, for example, the educational expenditure or the years of schooling, see Greiner *et al.* (2005, Ch. 4). Another approach is output oriented. In this context then the quality of school seems most important and a readily available measure of school quality are proficiency tests. The current study also uses an output-based measure and employs high school proficiency test passage rates because they reliably seem to predict labor market productivity and incomes (Sander, 1996; Loury and Garman, 1995; Murnane, Willett and Levy 1995). Crown and Wheat (1995) find that increases in education help explain the convergence of incomes in the U.S. South to other regions, further underscoring the link between education, income growth, and income distribution.

A large literature examines the convergence of incomes across regions. Kubo (1995) presents a theoretical model showing how regional development can be uneven, stable, or a mixture of uneven or stable across regions. The empirical literature seems to support all these scenarios. Webber, White and Allen (2005) find U.S. states' incomes are generally converging, although some states are converging more quickly than others. Choi (2004), in contrast, finds little evidence of overall output convergence across the U.S., but finds some convergence between neighboring states. In a similar vein, Bishop, Formby and Thistle (1994) find divergence in incomes during the 1970s and 1980s.

Partridge (2005) specifically examines the link between income distribution and growth. Partridge allows for both short-run and long-run responses of income distribution to growth, and also allows for separate effects of the tails and middle of the distribution. After making these adjustments, Partridge finds that the middle-class share of income is positively related to long-run growth, as is overall income inequality. Ohio, for example, in 1999 has a highly even distribution of incomes: it ranks 40th out of 50 states in income inequality with a Gini coefficient of 0.492 (Lynch, 2003). At the same time, Ohio had an above-average increase in inequality between 1988 and 1999. Between these years, incomes in Ohio rose 3.3% in real terms, including 7.1% for the top quintile of households (Lynch, 2003).

A long and contentious literature investigates the determinants of student achievement, generally measured as proficiency test scores. As aforementioned, in contrast to input-based, we stress output-based measures of human capital. A convenient starting point is a review of the literature by Hanushek (1986), which suggests that student achievement is generally related to parent and peer characteristics, but not to school-specific inputs. Recent literature confirms the importance of parent and peer characteristics. Student achievement is positively related to the presence of two-parent households (Bonesronning, 2004; Brasington, 2007, 1999), parent income levels (Dee, 1998; Driscoll, Halcoussis and Svorny, 2003; Dewey, Husted, and Kenny, 2000), and parent or community education levels (Brasington, 2007, 1999; Dee, 1998; Bonesronning, 2004; Driscoll, Halcoussis and Svorny, 2003; Dewey, Husted, and Kenny, 2000). Some research finds that the percent of students switching schools depresses achievement (Dewey, Husted, and Kenny, 2000), although other studies find less consistent results (Coates, 2003; Brasington, 1999). Some of the empirical literature find what our model predicts, namely, poverty seems to lower student achievement (Figlio and Stone, 2001; Brasington, 1999; Dee, 1998; Driscoll, Halcoussis and Svorny, 2003).

A school's competitive environment may also be related to student achievement. Studies find that private school market share is positively (Dee, 1998; Driscoll, Halcoussis and Svorny, 2003), negatively (Zanzig, 1997), or unrelated (Brasington, 2007) to public school performance. Competition from other public schools may also matter. The number of public school districts in a county has been found to increase student achievement (Figlio and Stone, 2001), sometimes increase it (Brasington, 2007), and increase it up to a certain point, then decrease it (Zanzig, 1997).

Although some of the recent literature still finds little relationship between school-specific inputs like teacher education levels and student achievement (Brasington, 1999, 2007; Coates, 2003; Bonesronning, 2004), other papers find a relationship. Student achievement has been found positively related to teacher salary (Sander, 1993; Zanzig, 1997; Dewey, Husted, and Kenny, 2000; Figlio, 1999), per-pupil expenditures (Dee, 1998; Bonesronning, 2004; Dewey, Husted, and Kenny, 2000), a low pupil to teacher ratio (Sander, 1993; Dewey, Husted, and Kenny, 2000; Figlio, 1999), and sometimes to teacher experience and education levels (Dewey, Husted, and Kenny, 2000).

3 The model

We next present a growth model that emphasizes the formation of human capital and the change in the social environment. We employ a Romer (1990) type of variety model that deals with two types of stock variables. One departure from the Romer type structure is that we introduce economies of agglomeration in the dynamics of social environment, which adds a strong nonlinearity to our model. With stronger agglomeration, the model possibly exhibits two basins of attraction and this allows us to explain the presence of persistent inequality.

3.1 Structure of the model

Consider a community with N_t homogeneous households. Two stock variables describe the state of a community at time t; the social environmental capital S_t and the human capital stock H_t. Different communities may have different S_t and H_t but all households in one community must be equally exposed to the community's H_t and S_t. We may think that communities are, by S_t and H_t, socioeconomically segregated. We assume that the number of households in a community remains constant over time.

A. Budget

We first assume that a community receives, let's say from the local government, a constant amount of transfer k per household, i.e., Nk in total, per unit of time.[3] The budget is specifically for the purpose of revitalizing the community either through investments i_t that create the environmental capital stock or through spending on immediate public services j_t that stimulates today's production. The transfer can not be carried over to the future and thus has to be exhausted at each time period. Then the allocation of the budget for a typical household in a community must satisfy

$$i_t + j_t = k. \tag{1}$$

Note that i_t creates the "stock" of environmental capital that can last for a while whereas j_t is a "flow" that only has a one-time effect on today's production or income.[4] We allow disinvestment of the environmental capital, $i_t \gtreqless 0$, while spending in production must be non-negative, $j_t \geq 0$. The divested portion can be used to increase j_t as long as (1) holds.

B. Community's social environment

The community's environmental stock S_t measures the general attractiveness of the community as a place for working, educating children, doing business, social contacts, etc.. It can change through the total investments Ni_t planned out of the budget k and/or through some private investments X_t that exogenously come from the outside world and add a variety of jobs and services to the community. We assume, for simplicity, that both types of investments equally contribute to the community's environment and depreciate at the same rate of δ. Then the law of motion of S_t is

$$\dot{S}_t = Ni_t + X_t - \delta S_t. \tag{2}$$

[3] We assume some wealth transfer mechanism through which some communities become recipients and others become payers of the transfer. In reality, the local government which conducts the transfer has some objective to attain and thus the transfer k should be endogenized. We, however, do not discuss this in our paper. We discuss only a recipient's welfare-maximizing problem assuming that k is exogenously given to the community.

[4] Investments i_t create the community's social environmental capital – examples: public facilities such as schools, parks, libraries, museums, transportation, sewers, security system, medical facilities that will last for a while. Spending on public services j_t contributes to the community's production that doesn't last – examples: patrolling policemen, park and street services, etc..

The size of exogenous investments X_t is, however, most likely to depend on the community's current environment S_t, i.e., the better the community's environment is, the more it can attract investments from the outside world and vice versa. Therefore, we define X_t to be a state-dependent variable

$$X_t = f(S_t). \tag{3}$$

This effect, $f' > 0$, which is based on some crowding-in effect, is an important feature of the so-called economies of agglomeration. This assumption makes the social environmental capital different from other types of capital such as physical capital. As it is unlikely that the scale of agglomeration increases with no limit, we additionally assume that it slows down after a certain point $\overline{S}$, i.e., $f'' > 0$ for $S_t < \overline{S}$, $f'' < 0$ for $S_t > \overline{S}$, and $\lim_{s_t \to \infty} f = S_{\max}$. Note that a public policy may indirectly spur, by increasing S_t through k, private investments X_t.

C. Production

The household's production y_t is a function of two input variables: various types of skilled labor $z(\theta)$ and the amount of spending on public services j_t. We assume that supply of unskilled labor is constant and common to all households and its contribution to production is reflected in the parameter γ. The parameter θ represents a type of skill development opportunity. Different types of skills and knowledge are acquired in different opportunities, e.g., represented by various types of schools and training programs and thus larger θ indicates that more diverse skill-development opportunities are provided to the household. We assume that better social environment provides more opportunities and thus the diversity is positively correlated to the per-household environmental capital S_t/N. Assuming that production follows the Cobb-Douglas type, production of a typical household in the community is

$$y_t = \gamma j_t^{\alpha} \int_0^{S_t/N} z(\theta)^{1-\alpha} d\theta \tag{4}$$

where different types of labor $z(\theta)$ are assumed to have the same efficiency $1-\alpha$ as in the Romer (1990) variety model for capital goods. This means that the quantity of labor of each type is the same $z(\theta) = \tilde{z}$ for all θ in equilibrium where the symbol $\widetilde{z}$ is the equilibrium size. Plugging this in (4) gives

$$y_t = \gamma j_t^{\alpha} \frac{S_t}{N} \tilde{z}^{1-\alpha}. \tag{5}$$

The total amount of labor used in production per household is $S_t\widetilde{z}/N$, so we define human capital per household h_t as[5]

$$\frac{S_t\widetilde{z}}{N} \equiv h_t. \tag{6}$$

Plugging this into (6) gives

$$y_t = \gamma j_t^{\alpha} s_t^{\alpha} h_t^{1-\alpha} \tag{7}$$

where $s_t \equiv S_t/N$. Notice that the community's environmental capital per household s_t is now explicitly in the production function.

D. Creation of human capital

Finally, the creation of human capital is achieved by forgone consumption

$$\dot{h}_t = \gamma j_t^{\alpha} s_t^{\alpha} h_t^{1-\alpha} - c_t. \tag{8}$$

F. Household's problem

The typical household maximizes the present value of the future utility stream. Let ρ denote the subjective discount rate. Then the household's problem is

$$\max_{\{c_t,\, i_t\}} \int_0^{\infty} u(c_t) e^{-\rho t} dt \tag{9}$$

subject to the budget constraint (1), the law of motion of h_t (8), and the law of motion of s_t

$$\dot{s}_t = i_t + N^{-1} f\left(S_t\right) - \delta s_t \tag{10}$$

for given initial states s_0 and h_0. The terminal condition is

$$\lim_{t\to\infty} \; s_t e^{-\rho t} = h_t e^{-\rho t} = 0. \tag{11}$$

3.2 Solving the problem

By adopting a CRRA-type utility function, the current-value Hamiltonian is

$$H_t = \frac{c_t^{1-\xi}}{1-\xi} + \lambda_t \left(i_t + N^{-1} f\left(S_t\right) - \delta s_t\right) + \mu_t \left(\gamma (k - i_t)^{\alpha} s_t^{\alpha} h_t^{1-\alpha} - c_t\right). \tag{12}$$

where ξ is the coefficient of relative risk aversion and $\xi > 0$.

[5] Note that in parallel to Romer's (1990) variety model where he defined the aggregate capital stock as $K_t \equiv \eta A_t \widetilde{x}$ with K_t as capital stock, A_t as stock of knowledge and $\widetilde{x}$ as marginal products of each capital good, we take η equal to one.

Necessary conditions

Applying the maximum principle, the optimal solution to (9) must satisfy the following conditions:

$$\mu_t = c_t^{-\xi}, \tag{13}$$

$$\lambda_t - \alpha\gamma\mu_t(k - i_t)^{\alpha-1} s_t^{\alpha} h_t^{1-\alpha} = 0, \tag{14}$$

$$\frac{\dot{\lambda}_t}{\lambda_t} = \rho + \delta - f'(S_t) - \frac{\mu_t}{\lambda_t}\alpha\gamma(k - i_t)^{\alpha} s_t^{\alpha-1} h_t^{1-\alpha}, \tag{15}$$

and

$$\frac{\dot{\mu}_t}{\mu_t} = \rho - (1-\alpha)\gamma(k - i_t)^{\alpha} s_t^{\alpha} h_t^{-\alpha}. \tag{16}$$

From a time derivative of (13) and (16), we obtain

$$\frac{\dot{c}_t}{c_t} = -\frac{1}{\xi}\frac{\dot{\mu}_t}{\mu_t}. \tag{17}$$

Rearranging (14) gives

$$\frac{\lambda_t}{\mu_t} = \alpha\gamma\left(k - i_t\right)^{\alpha-1} s_t^{\alpha} h_t^{1-\alpha}. \tag{18}$$

and inserting this into (15) gives

$$\frac{\dot{\lambda}}{\lambda} = \rho + \delta - f'(S_t) - \left(k - i_t\right) s_t^{-1}. \tag{19}$$

From (17) and a time derivative of log of (18), we get the law of motion of i_t as

$$\dot{i}_t = \frac{k - i_t}{1-\alpha}\left\{\frac{\dot{\lambda}_t}{\lambda_t} + \xi\frac{\dot{c}_t}{c_t} - \alpha\frac{\dot{s}_t}{s_t} - (1-\alpha)\frac{\dot{h}_t}{h_t}\right\}. \tag{20}$$

Steady States

At a steady state, $\dot{s}_t = \dot{h}_t = \dot{c}_t = \dot{i}_t = \dot{\lambda}_t = \dot{\mu}_t = 0$ should hold. Then, from (10),

$$i^* = \delta s^* - N^{-1} f\left(S^*\right) \tag{21}$$

and from (19),

$$j^* = k - i^* = \left[\rho + \delta - f'\left(S^*\right)\right] s^* \tag{22}$$

where an asterisk indicates the steady-state level of a variable and $S_t^* = Ns^*$

We may solve (21) and (22) for i^* and s^* and possibly find three steady states. This may happen due to economies of agglomeration assumed in the law of motion of s_t (10), i.e., the function f has local increasing returns to scale.[6]

[6] Note that if there is a constant positive growth in the number of households in a community, $\dot{N}_t/N_t \equiv g > 0$, then $\lim_{t\to\infty} N_t^{-1} f\left(S_t\right) = 0$ in (21) and $\lim_{t\to\infty} f'\left(S_t\right) = 0$ in (22). At a steady state, then, $i^* = \delta s^*$ and $k - i^* = \left(\rho + \delta\right) s^*$ hold. Solving these gives a unique steady state, $s^* = k/\left(\rho + 2\delta\right)$ and $i^* = \delta k/\left(\rho + 2\delta\right)$. This guarantees a non-explosive long-run growth of a community, i.e., $\dot{S}_t/S_t = \dot{H}_t/H_t = \dot{N}_t/N_t \equiv g$.

3.3 Numerical examples

It is unique to our model that economies of agglomeration $f(S_t)$ are introduced in the law of motion of s_t (10). Economies of agglomeration create a mechanism where households in a better environment enjoys a greater advantage in growth. Such a mechanism may exist clearly in some social settings but only weakly in others. Our next study numerically explores how the size of economies of agglomeration affects the dynamics of our model and the steady state values.

Let us use the following specific function:

$$f(S_t) = \frac{mS_t^\omega}{n^\omega + S_t^\omega} \tag{23}$$

where $m > 0$ determines the upper limit of f and $n^\omega > 0$ determines the size of economies of agglomeration for a given $\omega > 1$.[7] Note that a smaller n shifts f up everywhere and thus gives greater economies of agglomeration.

Steady states and local dynamics

Next, we want to give a rough sketch of the forces creating persistence of inequality. In order to do so, we vary the size of economies of agglomeration n^ω without changing the upper limit m. In this case, reducing n^ω has a bottom-up effect, i.e., a household with a smaller s_t tends to have a larger increase in the benefit from the agglomeration effect. Table 1 reports the steady state values for different sizes of economies of agglomeration $n = 100$, 95, 90, 85, and 80 for $\omega = 2$ and $m = 10$. It also reports the eigenvalues to each steady state, from which we can infer the local dynamics. Other parameters used are $\rho = .03$, $\xi = 2.5$, $k = 1$, $\delta = .05$, $\gamma = 1$, and $\alpha = .5$. The number of households is normalized as $N = 1$. We immediately observe the following: (1) With the weakest economies of agglomeration, $n = 100$, we have a unique steady state SS_1. By gradually reducing n to 95, 90, and 85, i.e., for stronger economies of agglomeration, two other upper steady states, SS_2 and SS_3, emerge. By further reducing n to 80, the two lower steady states, SS_1 and SS_2, disappear and the upper steady state SS_3 remains as a sole steady state. (2) From the eigenvalues reported in the last column of Table 1, we can see that the lower steady state SS_1 and the upper steady state SS_3 are always saddle stable points and the middle steady state SS_2 is always an unstable point. Thus, the middle steady state SS_2, if it exists, will never be reached in the long run unless the initial state of s_t and h_t happens to be equal to these steady state values. Either SS_1 or SS_3 can be reached from a given initial state. Some important economic implications can be derived from the observation.

[7] The parameter ω determines the curvature of the function $f(S_t)$. For $\omega > 1$, the function f has an s-shape, i.e., increaseing returns and then decreasing returns. For $0 < \omega < 1$, the function f is concave. For $\omega = 0$, f takes a constant value at $m/2$.

	y^*	i^*	j^*	s^*	h^*	eigenvalues
***n* = 100** (weaker agglomeration)						
SS_1	99.5	0.399	0.601	9.93	1658	0.117, -0.087, 0.038, -0.008
***n* = 95**						
SS_1	103.8	0.402	0.598	10.41	1729	0.110, -0.080, 0.038, -0.008
SS_2	463.6	0.398	0.602	46.18	7727	0.056, -0.026, 0.025, 0.005
SS_3	1161.4	-0.007	1.007	69.21	19357	0.066, -0.036, 0.033, -0.003
***n* = 90**						
SS_1	110.3	0.406	0.594	11.14	1838	0.100, -0.070, 0.038, -0.008
SS_2	302.8	0.458	0.542	33.52	5047	0.052, -0.022, 0.015+0.009i, 0.015-0.009i
SS_3	1844.1	-0.401	1.401	78.96	30735	0.076, -0.046, 0.035, -0.005
***n* = 85**						
SS_1	122.4	0.414	0.586	12.53	2041	0.085, -0.055, 0.037, -0.007
SS_2	222.0	0.457	0.543	24.56	3701	0.051, -0.021, 0.015+0.013i, 0.015-0.013i
SS_3	2449.4	-0.744	1.744	84.29	40824	0.083, -0.053, 0.036, -0.006
***n* = 80** (stronger agglomeration)						
SS_3	3038.2	-1.07	2.07	87.88	50637	0.091, -0.061, 0.037, -0.007

Table 1: Agglomeration effect n^ω on steady states when $\omega = 2$

Poverty or take-off – Local dynamics show that there are two basins of attraction in our model, one about the lower stable steady state SS_1 and the other about the upper stable steady state SS_3. We may call the lower basin a *poverty trap* and the upper basin a *take-off* region as the steady state values of income y^*[8], environmental capital stock s^*, and human capital stock h^* are all lower at SS_1 than they are at SS_3. There should be also a threshold separating the two basins. It leads to an open-ended dynamics where a community with the initial human and environmental resources above this threshold tends towards a better steady state SS_3 while a community with those below the threshold tends towards a poorer state SS_1. Note that local dynamic analysis as reported in Table 1, however, does not always help us to find the exact location of the threshold and we may not know which steady state is actually reached for a given initial state. Studying this issue requires us to undertake a global dynamic analysis, which we will discuss later.

Benefit from better environment – Under weaker economies of agglomeration, the benefit of improving the environment may not exceed its cost. In some cases, e.g., $n = 100$, all communities with any environmental capital stock tend towards the lower steady state SS_1 and maintain that lower level of environmental capital stock. For a stronger agglomeration effect, e.g., $n = 95$, 90, and 85, the benefit may exceed the cost for some communities and they take off for the upper steady state SS_3 while others still move to the lower steady state SS_1. The take-off happens to communities with the largest environmental capital stock first, then to more communities with less environmental capital stock under stronger economies of agglomeration, and eventually to all communities with any environmental capital stock, e.g., $n = 80$.

Investment in environment – The community in the upper steady state SS_3 invests less in the environment than the community in the lower steady state SS_1 even if it has more environmental capital to maintain. The reason that the environment at

[8]The steady state levels of consumption and income are equal as we assumed zero depreciation of human capital. This assumption, however, can be relaxed by introducing positive depreciation of human capital without changing the model's implications.

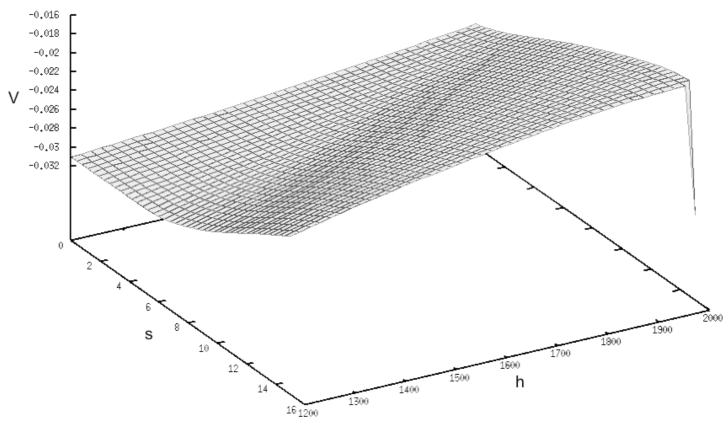

Figure 1: Value function $V(s_t, h_t)$ for $n = 100$

SS_3 can be maintained effortlessly, i.e., with less investment or even with divestment, is that its greater environment attracts bigger external investments based on some crowding-in or agglomeration effect. The community in the lower steady state, on the other hand, benefits less from this effect as it has a less attractive environment. This may explain a mechanism of persistent inequality.

Derivation of policy functions and out-of-steady-state dynamics

A policy function returns the best response to the current state. From the first order condition (13), the policy function for consumption is

$$c(s_t, h_t) = V_h(s_t, h_t)^{-\frac{1}{\xi}} \tag{24}$$

and from (14), the policy function for investment in environmental capital is

$$i(s_t, h_t) = k - \left[\alpha\gamma s_t^{\alpha} h_t^{1-\alpha} \frac{V_h(s_t, h_t)}{V_s(s_t, h_t)}\right]^{\frac{1}{1-\alpha}} \tag{25}$$

where the partial derivatives of the present value function, $V_s(s_t, h_t)$ and $V_h(s_t, h_t)$ are equal to their costate variables λ_t and μ_t respectively.The value function $V(s_t, h_t)$ can be numerically computed. Figure 1 shows the value function for $n = 100$, a unique steady state example in Table 1, computed by a numerical algorithm[9]. Not surprisingly, we find that the welfare is monotonically increasing in the environmental capital s_t and in the human capital h_t, $V_s > 0$ and $V_h > 0$, but the magnitude of their

[9] See Appendix 1 of how those two decision variables are numerically computed. We also want to note that the two decision variables, c_t and i_t, do not have to be chosen completely optimally. Their magnitude should only approximately be correct for the out of the steady state dynamics.

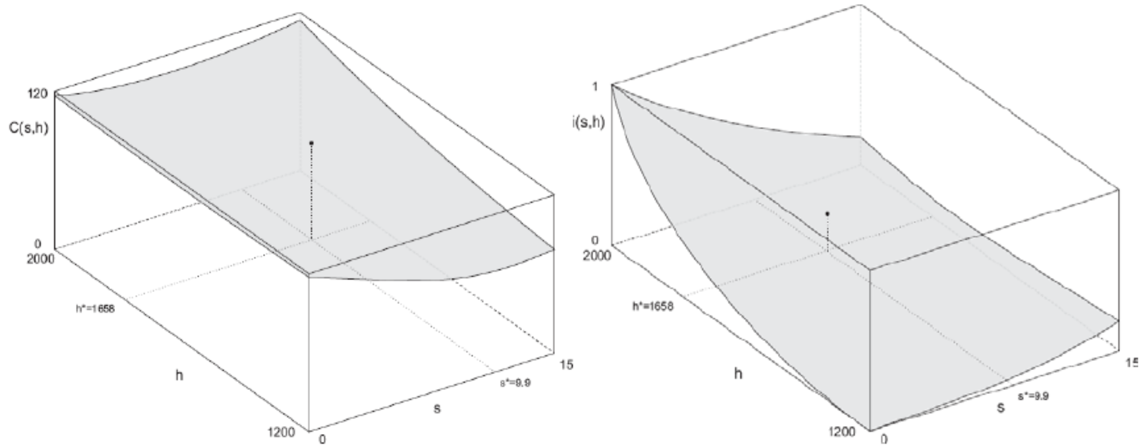

Figure 2: Policy functions $c\left(s_t, h_t\right)$ (left) and $i\left(s_t, h_t\right)$ (right) for $n = 100$

increase can be non-monotonic. Accordingly, the policy functions (24) and (25), that depend on the magnitude of V_s and V_h, can also be non-monotonic. Figure 2 shows the corresponding policy functions for $n = 100$.

We may summarize our numerical results as follows:

Remark 1: The welfare rises fastest from the origin to the south east direction, i.e., when the environmental capital and the human capital are both increasing (Figure 1). Welfare reaches the largest level for the largest s_t and h_t.

Remark 2: Consumption c_t monotonically increases as human capital h_t increases for a given environmental capital s_t (Figure 2, left). This may be so since human capital growth originates in foregone consumption (see eq. 8).

Remark 3: The effect of environmental capital s_t on consumption c_t for a given human capital h_t is non-monotonic (Figure 2, left). This is reasonable since consumption decision depends (out of the steady state) in a nonlinear way on the other policy variable.

Remark 4: The effect of human capital h_t on investment in environment i_t for a given environmental capital s_t is non-monotonic too (Figure 2, right). The investment i_t, however, may monotonically increase when human capital h_t increases along the steady state value of environmental capital s^*.

Remark 5: The effect of environmental capital s_t on investment in environment i_t for a given human capital h_t is also non-monotonic (Figure 2, right). The investment i_t, however, may monotonically decrease when environmental capital s_t increases along the steady state value of human capital h^*. This implies that a household in a richer environment may allocate a larger portion of the transfer to the one-time spending in production j_t.

Thus, overall though the policy variables, c_t and i_t, tend to move non-monotonically with the state variables (out of the steady state), they produce exactly what one expects for the welfare function: with higher s_t and h_t, welfare rises. When multiple steady states arise, deriving the shape of the value (welfare) function and the policy functions is not as easy as doing it for the example with a unique steady state. As there are two candidate paths associated with the two stable steady states, we are required to compare the present values of these two candidate paths for a given initial state and then find the global maximum of the value function. This is equivalent to undertaking a global dynamic analysis in a higher dimensional problem.

Recent technical development in numerical algorithms may help us to analyze the global dynamics of such a case. Grüne *et al.* (2005), for example, study a model with one control variable and one state variable, Haunschmied *et al.* (2003) and Grüne and Semmler (2004) study a model with one control variable and two state variables. As our model involves two control variables and two state variables, the dimension is even higher than for these existing studies. The numerical methods to solve this problem have not been developed far enough, but the existing studies let us make a conjecture on possible scenarios that may arise from our model. Three different scenarios are expected to appear in the following order as the size of economies of agglomeration increases, i.e., as n decreases:

Dominance of poverty – In this scenario, the path to the lower steady state SS_1 yields the largest welfare[10] for any given initial environmental and human resources. The path to the upper steady state SS_3, even if it exists, will never be reached by any community. Thus all communities will end up in a low steady state.

State-dependent dynamics[11] – A community will reach the lower steady state SS_1 or the upper steady state SS_3 depending on its initial environmental and human resources. There exists a threshold of environmental and human resources, beyond which the path to the upper steady state SS_3 yields the largest welfare and below which the path to the lower steady state SS_1 yields the largest welfare. When the initial state happens to be exactly on the threshold, the two paths yield the same welfare.

Dominance of take-off – In this scenario, the path to the upper steady state SS_3 yields the largest welfare for any given initial environmental and human resources. All communities necessarily take off and move to a high steady state.

Inequality and policy measures

The existence of a unique steady state guarantees a gradual convergence of income no matter how large the dispersion of initial status is. Inter-community inequality then tends to disappear as time passes. With multiple steady states, on the other hand,

[10] Here defined as welfare for the community.

[11] This case was discovered by Skiba (1978).

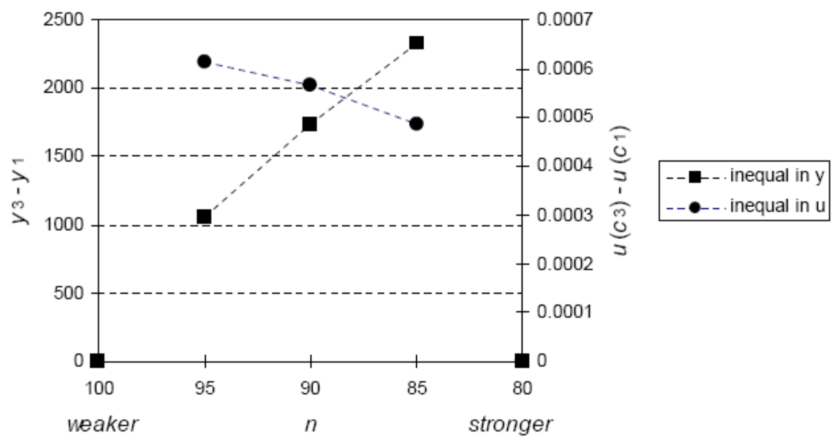

Figure 3: Inequality by size of agglomeration n^{ω} when $\omega = 2$

poorer communities tend towards the lower steady state SS_1while richer communities tend towards the upper steady state SS_3 and thus inter-community inequality may persist. It is true that the model's outcome, even if it shows potentially large inequality, is a result of a welfare-maximizing decision by households in a community; on the other hand the model also conveys that a series of decisions taken in a community is crucially dependent on the given initial human and environmental resources in a community. In reality, there are also a number of undesirable outcomes that are likely to arise from substantial inequality such as social instability and strong tensions between the poor and the rich. When policymakers are concerned with these issues, some policies to reduce inequality may be advisable.

Two measures can be used to describe the potential inter-community inequality – the size of inequality in terms of income and the size of inequality in terms of welfare. Income inequality can be measured by the distance between two income values at the two stable steady states, $y_3^* - y_1^*$ while inequality in welfare can be measure by the gap in welfare from two consumption values at the two stable steady states, $u(c_3^*) - u(c_1^*)$. Using these measures, Figure 3 reports inequalities computed from our numerical results in Table 1. There is no potential inequality with a unique steady state, e.g., $n = 100$ and 80. Otherwise greater inter-community inequality in income is inherent in the case where stronger economies of agglomeration exist, e.g., $n = 95$, 90, and 85. In those cases, however, inequality in terms of welfare can be smaller as Figure 3 shows. This can happen for the following reason: income and consumption increase at both steady states under stronger economies of agglomeration (Table 1) but the marginal utility from consumption for the poor at the lower steady state SS_1 is higher than that for the rich at the upper steady state SS_3 due to the concavity of utility function and therefore it may narrow inequality in welfare.

	y^*	i^*	j^*	s^*	h^*	eigenvalues
***k* = 1.00**						
SS_1	103.8	0.402	0.598	10.41	1729	0.110, -0.080, 0.038, -0.008
SS_2	463.6	0.398	0.602	46.18	7727	0.056, -0.026, 0.025, 0.005
SS_3	1161.4	-0.007	1.007	69.21	19357	0.066, -0.036, 0.033, -0.003
***k* = 1.01** (+1%)						
SS_1	106.3	0.406	0.604	10.57	1772	0.109, -0.079, 0.038, -0.008
SS_2	459.8	0.406	0.604	45.68	7662	0.055, -0.025, 0.024, 0.006
SS_3	1186.9	-0.013	1.023	69.62	19782	0.067, -0.037, 0.033, -0.003
***k* = 1.03** (+3%)						
SS_1	111.7	0.415	0.615	10.89	1861	0.108, -0.078, 0.038, -0.008
SS_2	452.9	0.422	0.608	44.71	7549	0.055, -0.025, 0.024, 0.006
SS_3	1237.0	-0.025	1.055	70.38	20616	0.067, -0.037, 0.033, -0.003
***k* = 1.05** (+5%)						
SS_1	117.2	0.424	0.626	11.23	1953	0.106, -0.076, 0.038, -0.008
SS_2	447.0	0.437	0.613	43.78	7450	0.055, -0.025, 0.023, 0.007
SS_3	1285.9	-0.035	1.085	71.09	21431	0.068, -0.038, 0.034, -0.004

Table 2: Effect of transfer k on steady states

Given various social and economic factors that determine the size of agglomeration n of an economy, we are interested here in investigating the effect of variation in the transfer k, as a policy parameter, on steady state values and inequality.[12] The transfer k in our model is given to a household as a device to stimulate investment in environmental capital and spending in production which is either consumed or invested in human capital. Galor and Zeira (1993), for example, show that subsidies that encourage investment in human capital reshape the economy's long-run distribution of wealth to reduce poverty. It is easy to predict that the steady state incomes increase as k rises while its effect on potential inequality is not so predictable. We first use the numerical example $n = 95$ in Table 1 as a benchmark case where the amount of transfer is set to $k = 1$ and then apply different higher transfer amounts, $k = 1.01$ (1% increase), 1.03 (+3%), and 1.05 (+5%), holding other parameters constant. The new steady state values and the eigenvalues are reported in Table 2 and potential inequalities are computed using the two different inequality measures in Figure 4. The economic implications are as follows:

Remark 1: The middle steady state SS_2 is always an unstable one, see the corresponding eigenvalues (Table 2). Only the lower steady state SS_1 or the upper steady state SS_3 can be reached in the long run. Thus potential inequality remains positive for all ks in Table 2.

Remark 2: When a larger transfer amount is provided equally to each household, the values of income y^*, environmental capital stock s^*, and human capital stock h^* at both stable steady states, SS_1 and SS_3, increase (Table 2).

Remark 3: Although the economic status of communities at both stable steady states improve by a larger transfer k, Figure 4 shows that the same policy may enlarge potential income inequality. The intuitive explanation for this is that marginal returns

[12] We do not formally discuss a policy objective of varying k here but simply want to indicate the direction to which a community possibly moves by varying k.

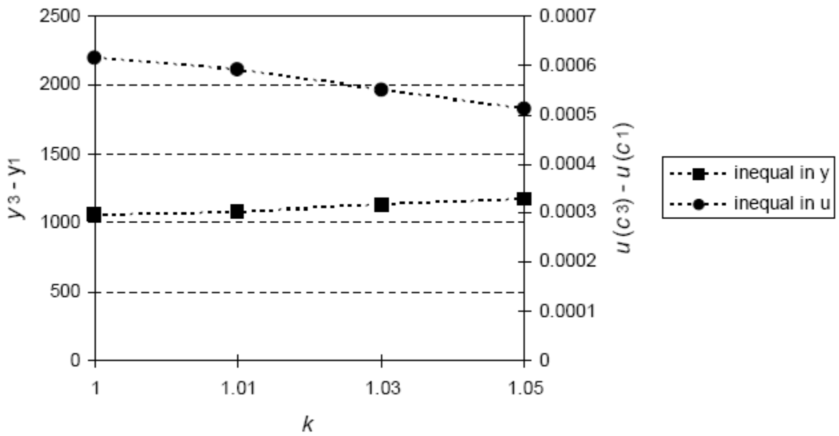

Figure 4: Inequality by amount of transfer k

are higher by investing an additional dollar in a high-income community at SS_3 than investing it in a low-income community at SS_1.

Remark 4: An increase in income inequality, however, does not immediately mean an increase in inequality in terms of welfare. The impact of the transfer on inequality on welfare can be opposite to its impact on income inequality as Figure 4 exemplifies. This is because a low-income community at SS_1 has higher marginal utility than a high-income community at SS_3 under a concave utility function.

Increasing k equally to high- and low-income households may not reduce income inequality. It may, however, reduce inequality in welfare. Public policy aimed at reducing inequality is justified in part because greater inequality often creates greater social tensions and instability. So although public policy may not reduce income inequality, it may still help reduce social instability by reducing welfare inequality.

Remark 5: Thus, as far as such negative externalities are concerned, the highest transfer should be given to the region which has the largest inter-community welfare gap.

4 Poverty trap in data

The model predicts that human capital and social environment are positively correlated across steady states and so are income and consumption. Therefore, some communities should exhibit high levels of all these variables – income, consumption, human capital and social environment whereas in other communities the levels of all these four variables should be low. To partially test for whether the four variables are positively correlated across communities, one may pick out two of these variables and study the relationship between them. For a measure of human capital, the use of school performance or student achievement, as we discussed in Subsection 2.2, is often justified. Concerning a measure of social environment, there are many attributes that make a community more or less desirable. Such attributes include the level of crime, pollution, and the availability of parks and cultural institutions like museums and libraries. Such attributes are reflected in the value of houses. Holding the physical characteristics of a house constant, a house in an environment with less crime, less pollution, more parks and museums will be worth more than a house without such a desirable social environment (Clark and Kahn, 1988).

Income and human capital – Poverty is indeed routinely found to depress student achievement in education production function studies (Hanushek, 1986). For example, Callan and Santerre (1990) find that the percent of students' families receiving welfare is negatively related to proficiency test scores in Connecticut, and that it is also linked to a lower percent of high school graduates pursuing higher education. Brasington (1999) finds poverty rates lower student proficiency test passage rates in Ohio, controlling for parent education levels. The elasticity of math proficiency passage rates with respect to poverty is -0.06, implying that a ten percent rise in poverty would lower proficiency passage by about half a percent.

Social environment and human capital – A great deal of literature finds a link between social environment and human capital, as measured by house prices and public school quality. Much of the older school quality capitalization literature uses spending per pupil to measure school quality (e.g. Oates, 1969), but school quality has been measured by the pupil-teacher ratio (Hoehn, *et al.* 1987), a state-assigned school competitiveness index (Hite *et al.*, 2001), state-assigned grades to schools (Figlio and Lucas, 2004), and the value-added of schooling (Downes and Zabel, 2002; Hayes and Taylor, 1996). Many studies find a correlation between house prices and proficiency test scores (e.g. Haurin and Brasington, 1996; Jud and Watts, 1981; Hite, *et al.* 2001; Hayes and Taylor, 1996; Figlio and Lucas, 2004). An extensive review of the literature on the correlation between house prices and school quality is found in Ross and Yinger (1999).

The model also predicts a state-dependent transition of these variable in out-of-steady-state dynamics. Here we pick out human capital and test whether the transition of educational attainment as a measure of human capital is affected by the past state of attainment. The data we use cover math proficiency passage rates of 608 school districts in Ohio during 1990-2002. Ohio is about as representative of the U.S. as any state gets. It has six fairly large urban areas with population between 600,000 and 2.2 million, along with numerous small cities and rural areas. It has prosperous suburban school districts and poverty-stricken inner cities, prosperous farming communities throughout the western and central parts of the state and poor Appalachian areas in the southeast. The uneven prosperity of the state led to a successful challenge of Ohio's school funding formula in the 1990s. In response, the state legislature increased tax revenue devoted to schools, with property-rich school districts getting

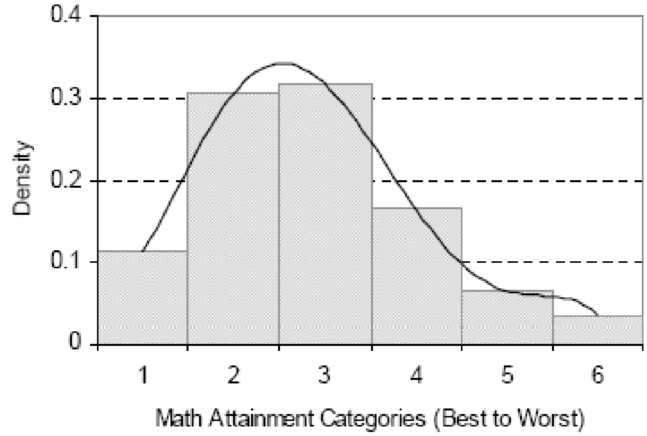

Figure 5: Ergodic distribution of math attainment

less state funding and property-poor school districts getting increased state funding. As a result, between the 1994-1995 school year and the 2005-2006 school year, nominal expenditures per pupil rose from $8188 to $13,558 in the relatively prosperous Cincinnati suburban school district of Princeton, and from $4204 to $9286 in the poor southeastern school district of Vinton County, representing 27% and 70% increases in real spending.

$t+1$ \ t	1	2	3	4	5	6	Ergodic dist.
1	0.57	0.1457	0.0114	0.0023	0.0026	0	0.1133
2	0.3833	0.5742	0.246	0.05	0.0087	0.0032	0.306
3	0.0467	0.2429	0.5413	0.3513	0.1046	0.0297	0.3167
4	0	0.0358	0.1708	0.44	0.3743	0.0871	0.1648
5	0	0.0014	0.0247	0.136	0.3928	0.2509	0.0647
6	0	0	0.0058	0.0204	0.117	0.6292	0.0344

Table 3: Transition matrix of math attainment

We break the percentage passage rates into 6 categories; category 1 (90-99% passage), category 2 (80-89% passage), category 3 (70-79% passage), category 4 (60-69% passage), category 5 (50-59% passage), category 6 (<50% passage), where category 1 is the best and category 6 is the worst, and then construct ten independent one-year transition matrices.[13] These are reported in Table 4 in Appendix 2. By simply taking the average of ten transition probabilities for each entry, we get a new transition matrix as shown in Table 3. It is a column stochastic matrix where its column sums are unity. The last column reports the resulting ergodic distribution assuming that the Markov chain has stationary transition probabilities (homogeneous chains).[14] Figure

[13] There is no data for 1994. It is excluded due to the major test procedure change by the Ohio Department of Education. Thus, both 1993-94 and 1994-95 transition matrices cannot be constructed. For the same reason, the transition from 1993 (before the change) to 1995 (after) is inappropriate for inclusion to derive the average transition.

[14] For details, see Appendix 2.

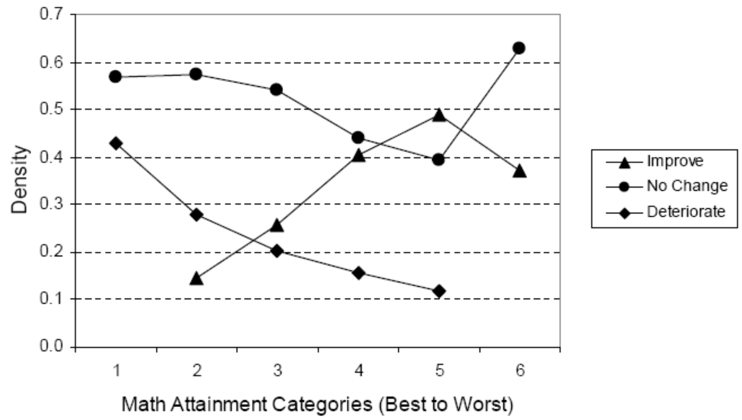

Figure 6: Transition of math attainment

5 shows the ergodic distribution of math attainment reported in Table 3. The ergodic distribution is the unique stationary distribution for a given trend of transition. The distribution is highly skewed to the right and it suggests no tendency to be normal in the long run. The majority of the school districts are highly concentrated about categories 2 and 3; on the other hand, there is no indication that school districts left behind in the worst two categories 5 and 6 disappear in the long run. This again implies the existence of two domains of attraction in the transition mechanism; the lower-attainment domain as an educational-poverty trap prevents some schools to take off for better categories and therefore inequality persists.

When we focus on the direction (improve, no change, or deteriorate) only of the change in math attainment, the transition matrix in Table 3 can be further summarized as Figure 6. For example, the chance that a school district that is in category 3 moves to any better category, i.e., categories 1 or 2, within a year is 25.74%, the chance that it stays in the same category is 54.13%, and the chance that it moves to any worse category, i.e., categories 4, 5, or 6, is 20.13%. Although both categories 5 and 6 are understood as low-performance categories, category 5 has the highest chance of take-off (49%) and the lowest chance of immobility (39.28%) while category 6 has the highest chance of immobility (62.92%) and a much lower chance of take-off (37.08%). This indicates that there may be a threshold between categories 5 and 6 separating a take-off region and a poverty trap.

5 Conclusions

This paper analyzes the role of environmental and human capital formation in explaining the presence of inter-community inequality that persist. We study, in a Romer (1990) type variety model, a welfare-maximizing problem undertaken in a community that is characterized by environmental capital and human capital. We incorporate the so-called economies of agglomeration (or crowding-in effect) in the underlying dynamics of social environment. With a larger size of economies of agglomeration, our model exhibits two basins of attraction: one basin interpreted as a poverty trap and the other basin interpreted as a take-off region. Whenever the initial environmental and human resources fall within the poverty trap (within the take-off region), a community tends towards the lower steady state (towards the upper steady state) where income, consumption, human capital, and environmental capital are all lower (higher) than those at the other steady state. Such global dynamics give rise to state-dependent dynamics and data on educational attainment as a measure of human capital, for example, seem to support this point.

With the help of a numerical method, we derive the value function and the policy functions that can be a guidance to a community's decision on consumption and budget allocation between investments in environment and production out of the steady state. Communities benefit from economies of agglomeration. Stronger economies of agglomeration bring about greater environmental and human resources and higher income in both high- and low-income communities. Our numerical study, however, suggests that income inequality between the high- and low-income communities may rise with stronger economies of agglomeration while inequality measured by welfare may decline as marginal utility slopes down.

Finally, we study a policy that aims at reducing inter-community inequality. We investigate the effect of a policy of transfers on income and welfare inequality. We find that larger transfers may increase income inequality while reducing welfare inequality, because of decreasing marginal utility. Inequality in welfare may be a more appealing measure when a policy intends to reduce unwanted externalities from inequality (such as crimes). In that case, a policy recommendation may be that the highest transfer should be given to the region which has the largest inter-community welfare gap.

References

[1] Arrow, Kenneth J. 1963. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29(3): 155-173.

[2] Becker, Gary S. and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87(6): 1153-1189.

[3] Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life of Earnings." *Journal of Political Economy* 75(4): 352-365.

[4] Bishop, John A., John P. Formby, and Paul D. Thistle. 1994. "Convergence and Divergence of Regional Income Distributions and Welfare." *Review of Economics and Statistics* 76(2): 228-235.

[5] Bonesronning, Hans. 2004. "Can Effective Teacher Behavior Be Identified?" *Economics of Education Review* 23(3): 237-247.

[6] Bowles, Samuel and Herbert Gintis. 2002. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16(3): 3-30.

[7] Brasington, David M. 1999. "Central City School Administrative Policy: Systematically Passing Undeserving Students." *Economics of Education Review* 18(2): 201-212.

[8] Brasington, David M. 2007. "Public and Private School Competition: The Spatial Education Production Function" in T. Asada and T. Ishikawa (eds.) *Time and Space in Economics*, Springer-Verlag.

[9] Brock, A. William and Steven N. Durlauf. 2006. "Social Interactions and Macroeconomics" in D. Colander (ed.) *Post-Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge University Press.

[10] Callan, Scott J. and Rexford E. Santerre. 1990. "The Production Characteristics of Local Public Education: A Multiple Product and Input Analysis." *Southern Economics Journal* 57: 468-480.

[11] Choi, Chi-Young. 2004. "A Reexamination of Output Convergence in the U.S. States: Toward Which Level(s) Are They Converging?" *Journal of Regional Science* 44(4): 713-741.

[12] Clark, David E. and James R. Kahn. 1988. "The Social Benefits of Urban Cultural Amenities." *Journal of Regional Science* 28(3): 363-377.

[13] Coates, Dennis. 2003. "Education Production Functions Using Instructional Time as an Input." *Education Economics* 11(3): 273-292.

[14] Crown, William H. and Leonard F. Wheat. 1995. "State Per Capita Income Convergence Since 1950: Sharecropping's Demise and Other Influences." *Journal of Regional Science* 35(4): 527-552.

[15] Dee, Thomas S. 1998. "Competition and the Quality of Public Schools." *Economics of Education Review* 17(4): 419-427.

[16] Dewey, James, Thomas A. Husted and Lawrence W. Kenny. 2000. "The Ineffectiveness of School Inputs: A Product of Misspecification?" *Economics of Education Review* 19(1): 27-45.

[17] Downes, Thomas A. and Jeffrey E. Zabel. 2002. "The Impact of School Characteristics on House Prices: Chicago 1987-1991" *Journal of Urban Economics* 52: 1-25.

[18] Driscoll, Donna, Dennis Halcoussis and Shirley Svorny. 2003. "School District Size and Student Performance." *Economics of Education Review* 22(2): 193-201.

[19] Durlauf, Steven N. 2004. "Neighborhood Effects" in J. V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, vol. 4, North Holland.

[20] Durlauf, Steven N. 2006. "Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps" in S. Bowles, S. Durlauf, and K. Hoff (eds.) *Poverty Traps*. Princeton University Press.

[21] Figlio, David N. 1999. "Functional Form and the Estimated Effects of School Resources." *Economics of Education Review* 18(2): 241-252.

[22] Figlio, David N. and Maurice E. Lucas. 2004. "What's In a Grade? School Report Cards and the Housing Market." *American Economic Review* 94: 591-604.

[23] Figlio, David N. and Joe A. Stone. 2001. "Can Public Policy Affect Private School Cream Skimming?" *Journal of Urban Economics* 49(2): 240-266.

[24] Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60 (1): 35-52.

[25] Greiner, Alfred, Willi Semmler, and Gang Gong. 2005. *The Forces of Economic Growth: A Time Series Perspective*. Princeton University Press.

[26] Grüne, Lars, Mika Kato, and Willi Semmler. 2005. "Solving Ecological Management Problems Using Dynamic Programming." *Journal of Economic Behavior and Organization* 57: 448–473.

[27] Grüne, Lars and Willi Semmler. 2004. "Using Dynamic Programming with Adaptive Grid Scheme for Optimal Control Problems in Economics." *Journal of Economic Dynamics and Control* 28: 2427-2456.

[28] Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3): 1141-1177.

[29] Haunschmied, Josef L., Peter M. Kort, Richard F. Hartl, and Gustav Feichtinger. 2003. "A DNS-curve in a Two-State Capital Accumulation Model: a Numerical Analysis." *Journal of Economic Dynamics and Control* 27(4): 701-716.

[30] Haurin, Donald R. and David M. Brasington. 1996. "The Impact of School Quality on Real House Prices: Interjurisdictional Effects." *Journal of Housing Economics* 5(4): 351-368.

[31] Hayes, Kathy J. and Lori L. Taylor. 1996. "Neighborhood School Characteristics: What Signals Quality to Homebuyers?" *Federal Reserve Bank of Dallas Economic Review* 3: 2-9.

[32] Hite, Diane, Wen Chern, Fred Hitzhusen, and Alan Randall. 2001. "Property Value Impacts of an Environmental Disamenity: the Case of Landfills." *Journal of Real Estate Finance and Economics* 22: 185-202.

[33] Hoehn, John P., Mark C. Berger and Glenn C. Blomquist. 1987. "A Hedonic Model of Interregional Wages, Rents, and Amenity Values." *Journal of Regional Science* 27(4): 605-620.

[34] Jud, G. Donald and James M. Watts. 1981. "Schools and Housing Values." *Land Economics* 57: 459-470.

[35] Katz, Lawrence F. and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality" in Orley Ashtenfelter and David Card (eds.) *Handbook of Labor Economics*, vol. III, Elsevier.

[36] Kubo, Yuji. 1995. "Scale Economies, Regional Externalities, and the Possibility of Uneven Regional Development." *Journal of Regional Science* 35(1): 29-42.

[37] Loury, Linda D. and David Garman. 1995. "College Selectivity and Earnings." *Journal of Labor Economics* 13(2): 289-308.

[38] Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22: 3-42.

[39] Lynch, Robert G. 2003. "Estimates of Income and Income Inequality in the United States and in Each of the Fifty States: 1988-1999." *Journal of Regional Science* 43(3): 571-587.

[40] Matsuyama, Kiminori. 2000. "Endogenous Inequality.," *Review of Economic Studies*, 67(4): 743-759.

[41] Matsuyama, Kiminori. 2007. "Aggregate Implications of Credit Market Imperfections." *NBER* Working Paper, No. 13209.

[42] Mincer, Jacob. 1958. "Investment in Human Capital and Personal Distribution." *The Journal of Political Economy* 66(4): 281-302.

[43] Murnane, Richard J., John B. Willett and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77(2): 251-266.

[44] Oates, W. E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77: 957-971.

[45] Partridge, Mark D. 2005. "Does Income Distribution Affect U.S. State Economic Growth?" *Journal of Regional Science* 45(2): 363-394.

[46] Romer, Paul. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94(5): 1002-1037.

[47] Romer, Paul. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5): 71-102.

[48] Ross, Stephen L. and John Yinger. 1999. "Sorting and Voting: A Review of the Literature on Urban Public Finance." in P. Cheshire and E. S. Mills (eds.), *Handbook of Regional and Urban Economics*, North-Holland.

[49] Sander, William. 1993. "Expenditures and Student Achievement in Illinois: New Evidence." *Journal of Public Economics* 52(3): 403-416.

[50] Sander, William. 1996. "Catholic Grade Schools and Academic Achievement." *Journal of Human Resources* 31(3): 540-548.

[51] Semmler, Willi and Marvin Ofori. 2007. "On Poverty Traps, Thresholds and Take-Offs." *Structural Change and Economic Dynamics* 18: 1-26.

[52] Seneta, Eugene. 2006. *Non-negative Matrices and Markov Chains*. 2nd ed., Springer.

[53] Skiba, A. K. 1978. "Optimal Growth with a Convex-Concave Production Function." *Econometrica* 46: 527-539.

[54] Smith, Adam. 1852. *An Inquiry into the Nature and Causes of the Wealth of Nations*. T. Nelson and Sons, London, first edition 1776.

[55] Uzawa, Hirofumi. 1965. "Optimum Technical Change in an Aggregative Model of Economic Growth." *International Economic Review* 6(1): 18-31.

[56] Webber, Don J., Paul White, and David O. Allen. 2005. "Income Convergence across U.S. States: An Analysis Using Measures of Concordance and Discordance." *Journal of Regional Science* 45(3): 565-589.

[57] Zanzig, Blair R. 1997. "Measuring the Impact of Competition in Local Government Education Markets on the Cognitive Achievement of Students." *Economics of Education Review* 16(4): 431-444.

Appendix 1: Numerical solution method

We here briefly describe the dynamic programming algorithm as applied in Grüne and Semmler (2004) that enables us to numerically solve the dynamic model as proposed in Section 4. The feature of the dynamic programming algorithm is an adaptive discretization of the state space which leads to high numerical accuracy with moderate use of memory.

Such algorithm is applied to discounted infinite horizon optimal control problems of the type introduced in Section 4. In our model variants we have to numerically compute $V(x)$ for

$$V(x) = \max_u \int_0^\infty e^{-r} f(x,u)dt$$

$$s.t. \dot{x} = g(x,u)$$

where u represents the control variable and x a vector of state variables.

In the first step, the continuous time optimal control problem has to be replaced by a first order discrete time approximation given by

$$V_h(x) = \max_j J_h(x,u), \quad J_h(x,u) = h \sum_{i=0}^{\infty} (1-\theta h) U f(x_h(i), u_i) \tag{A1}$$

where x_u is defined by the discrete dynamics

$$x_h(0) = x, \; x_h(i+1) = x_h(i) + hg(x_i, u_i) \tag{A2}$$

and $h > 0$ is the discretization time step. Note that $j = (j_i)_{i \in \mathbb{N}_0}$ here denotes a discrete control sequence.

The optimal value function is the unique solution of a discrete Hamilton-Jacobi-Bellman equation such as

$$V_h(x) = \max_j \{hf(x,u_o) + (1+\theta h)V_h(x_h(1))\} \tag{A3}$$

where $x_h(1)$ denotes the discrete solution corresponding to the control and initial value x after one time step h. Abbreviating

$$T_h(V_h)(x) = \max_j \{hf(x,u_o) + (1-\theta h)V_h(x_h(1))\}, \tag{A4}$$

the second step of the algorithm now approximates the solution on grid Γ covering a compact subset of the state space, i.e. a compact interval $[0,K]$ in our setup. Denoting the nodes of Γ by $x^i, i = 1, ..., P$, we are now looking for an approximation V_h^Γ satisfying

$$V_h^\Gamma(X^i) = T_h(V_h^\Gamma)(X^i) \tag{A5}$$

for each node x^i of the grid, where the value of V_h^Γ for points x which are not grid points (these are needed for the evaluation of T_h) is determined by linear interpolation.

We refer to the paper cited above for the description of iterative methods for the solution of (A5). Note that an approximately optimal control law (in feedback form for the discrete dynamics) can be obtained from this approximation by taking the value $j^*(x) = j$ for j realizing the maximum in (A3), where V_h is replaced by V_h^Γ. This procedure in particular allows the numerical computation of approximately optimal trajectories.

In order the distribute the nodes of the grid efficiently, we make use of a posteriori error estimation. For each cell C_l of the grid Γ we compute

$$\eta_l := \max_{k \in c_l} \mid T_h(V_h^\Gamma)(k) - V_h^\Gamma(k) \mid .$$

More precisely we approximate this value by evaluating the right hand side in a number of test points. It can be shown that the error estimators η_l give upper and lower bounds for the real error (i.e., the difference between V_j and V_h^Γ) and hence serve as an indicator for a possible local refinement of the grid Γ. It should be noted that this adaptive refinement of the grid is very effective for computing steep value functions and models with multiple equilibria, see Grüne and Semmler (2004).

Code for the numerical solution method

This is a Numerical Program operating in Linux. The background is the Software Package developed by Lars Grüne, Bayreuth University Germany. Details of the Dynamic Programming algorithm used here can be found on the website of Lars Grüne.[15]

```
#ifdef _PROTOTYPES_
real povert_g( real *, real * );
void povert_f( real *, real *, real * );
#endif
#ifdef _DECLARATIONS_
{
 "poverty",            /* name */
 "Zur Erklärung der Modulerstellung",      /* help text */
 povert_g,                    /* cost function */
 povert_f,                    /* (projected) right hand side */
 povert_f,                    /* (unprojected) right hand side */
 dummy2,                      /* projection #1 (for calculation) */
 dummy2,                      /* projection #2 (for output) */
 dummy2,                      /* injection #1 */
 dummy2,                      /* injection #2 */
 2,                           /* system dimension */
 2,                           /* output dimension */
```

[15] http://www.math.uni-bayreuth.de/~lgruene/

```
 2,                              /* unprojected dimension */
 0.0,                              /* projection information */
 0,                              /* continuous/discrete time (0/1) */
 2,                              /* control dimension */
 { /* equation parameter specification */
 11, /* number of parameters */
 "Parameters",                   /* window title */
 { "em", "en ", "omeg ", "rho ", "zeta ", "k ", "delt ", "gam ", "alf ", "bet ", "el "},
 /* names of parameters */
 { 0.5, 65, 3, 0.03, 2.5, 1, 0.05, 1, 0.3, 0.3, 1 }, /* default values */
 "Parameters", /* help text */
 #ifdef __MKO__ /* widgets (do not change!) */
 NULL,
 NULL
 #endif
 },
},
#endif
#ifdef _IMPLEMENTATION_
#define em (equdata.paradata[equdata.act]->value[0])
#define en (equdata.paradata[equdata.act]->value[1])
#define omeg (equdata.paradata[equdata.act]->value[2])
#define rho (equdata.paradata[equdata.act]->value[3])
#define zeta (equdata.paradata[equdata.act]->value[4])
#define k (equdata.paradata[equdata.act]->value[5])
#define delt (equdata.paradata[equdata.act]->value[6])
#define gam (equdata.paradata[equdata.act]->value[7])
#define alf (equdata.paradata[equdata.act]->value[8])
#define bet (equdata.paradata[equdata.act]->value[9])
#define el (equdata.paradata[equdata.act]->value[10])
void povert_f( real *x, real *u, real *y )
{
real fs, u0,u1;
if (u[0]>1) u0=0.99;
if (u[0]<0) u0=0.01;
else u0=u[0];
if(u[1]<0) u1=0.01;
else u1=u[1];
fs=(em*pow(x[0],omeg))/(pow(en,omeg)+pow(x[0],omeg));
 y[0] = u0-delt*x[0]+fs;
 y[1] = gam*(pow((k-u0),alf))*pow(x[0],(alf+bet))*pow(el,bet)*pow(x[1],(1-alf-bet))-u1;
}
```

```
real povert_g( real *x, real *u )
{
real u1;
if(u[1]<0) u1=0.01;
else u1=u[1];
 return(pow(u1,(1-zeta))/(1-zeta));
}
#undef em
#undef en
#undef omeg
#undef rho
#undef zeta
#undef k
#undef delt
#undef gam
#undef alf
#undef bet
#undef el
#endif
```

Appendix 2: Derivation of ergodic distribution

Our study is based on 14 years of data (1990-2002) of math proficiency passage rates of 608 school districts in Ohio. We break the percentage passage rates into 6 categories; category 1 (90-99% passage), category 2 (80-89% passage), category 3 (70-79% passage), category 4 (60-69% passage), category 5 (50-59% passage), category 6 (<50% passage) so that for each of 14 years the number of school districts falls into each of the six categories. Category 1 is the best and category 6 is the worst. Note that there is no data for 1994. It is excluded due to the major test procedure change by the Ohio Department of Education. In 1994, Ohio started letting students take the 9th grade proficiency test in 8th grade. So the numbers reported for 9th grade proficiency include the 8th graders who passed.

Based on the data, we can construct 10 independent one-year transition matrices such as 1990-91, 91-92, ..., 2001-02. Those are reported in Table 4. Note that the transition matrix 1993-95 that crosses over the 1994 major change of the test procedure is discarded for the reason that it can be largely affected by the change. Each matrix is a column stochastic matrix where its column sums are unity. Then by simply taking the average transition probabilities, we obtain the averaged transition matrix for the 10 transition matrices. This is reported in Table 3 in Section 4. The percentage ergodic distribution of 608 school districts is reported in the last column of each transition matrix assuming that the Markov chain has stationary transition probabilities (homogeneous chains). The ergodic distribution is the unique stationary distribution and useful to estimate a long-run outcome based on the recent trend in the school performance. We find that most distributions are highly skewed to the right. Only a few school districts are left in categories 5 or 6 and the majority of the school districts end in categories 2 or 3.

We next summarize some important characteristics of an ergodic distribution. Let's P be the irreducible transition matrix. Then the postmultiplication by $\mathbf{1}$ the vector with unity in each position gives

$$\mathbf{1}'P = \mathbf{1}'$$

by stochasticity of P where 1 is an eigenvalue and $\mathbf{1}$ is a corresponding left eigenvector.

Since all column sums of P are equal and the Perron-Frobenius eigenvalue lies between the largest and the smallest

$$\min_j \sum_{i=1}^{n} p_{ij} \leq r \leq \max_j \sum_{i=1}^{n} p_{ij}$$

where $r \geq |\lambda|$ for any eigenvalue λ of P, 1 is the Perron-Frobenius eigenvalue of P and $\mathbf{1}$ is the corresponding left Perron-Frobenius eigenvector.

Let's define the corresponding right eigenvector as a column vector v that is normed as

$$\mathbf{1}'v = 1.$$

Then, we have

$$Pv = v$$

where v is the vector of probability distribution.

Theorem 1 *An irreducible Markov chain has a unique stationary distribution given by the solution v of $Pv = v$, $\mathbf{1}'v = 1$.*

Proof. Any initial probability distribution Π_0 is called a stationary distribution if

$$\Pi_0 = \Pi_k,\ k = 1, 2, ...$$

If Π_0 is a stationary distribution,

$$P\Pi_0 = \Pi_0,\ \Pi_0 \geq 0,\ \mathbf{1}'\Pi_0 = 1.$$

By uniqueness of the right Perron-Frobenius eigenvalue of P, $\Pi_0 = v$. ■

Theorem 2 *(Ergodic Theorem for primitive Markov chains) As $k \to \infty$, for a primitive Markov chain, $P^k \to v\mathbf{1}'$ elementwise where v is the unique stationary distribution of the Markov chain and the rate of approach to the limit is geometric.*

Proof. See Seneta (2006) Theorem 1.2, p. 9. ■

Corollary 3 *The unique stationary distribution is independent from the initial distribution.*

Proof. For any initial probability distribution Π_0, as $k \to \infty$

$$P^k \Pi_0 \to v\mathbf{1}'\Pi_0.$$

Since $\mathbf{1}'\Pi_0 = 1$,

$$P^k \Pi_0 \to v.$$

■

Code for the derivation of ergodic distribution

We used *Mathematica* to obtain the ergodic distributions reported in Table 4. For each 6×6 transition matrix, we first compute the eigenvalues, then find the corresponding right eigenvector to the Perron-Frobenius eigenvalue 1, and normalize it so that $\mathbf{1}'v = 1$ By using the following steps, the same results should be reproduced:

Step 1: Specify the transition matrix P

In[1]:= P = {{p11, p12, ..., p16}, {p21, p22, ..., p26},...{p61, p62, ..., p66}}; MatrixForm[P]

Step 2: Obtain eigenvalues

In[2]:= Eigenvalues[P]

Mathematica sorts eigenvalues, if they are numeric, in order of decreasing absolute value. Since the Perron-Frobenius eigenvalue of P is $r \geq |\lambda|$ for any eigenvalue λ of P, the first eigenvalue should be 1.

Step 3: Obtain eigenvectors

In[3]:= MatrixForm[Eigenvectors[P]]

Mathematica returns the matrix of eigenvectors. The corresponding Perron-Frobenius right eigenvector to the first eigenvalue 1 should be the first row.

Step 4: Normalize the eigenvector so that $\mathbf{1}'v = 1$.

The obtained v is the stationary distribution reported in the last column of each matrix in Table 4.

91\90	1	2	3	4	5	6	Ergodic dist
1	0.5	0.2667	0.0244	0	0	0	0.2154
2	0.5	0.5333	0.3171	0.0822	0.0068	0	0.3857
3	0	0.1333	0.5122	0.3425	0.1293	0.0121	0.1979
4	0	0.0667	0.122	0.3699	0.3401	0.1273	0.1127
5	0	0	0	0.137	0.4082	0.2879	0.0465
6	0	0	0.0244	0.0685	0.1156	0.5727	0.0419

92\91	1	2	3	4	5	6	Ergodic dist
1	0.5	0.1034	0	0	0	0	0.0388
2	0.3333	0.5517	0.1972	0.04	0.0061	0.0047	0.1873
3	0.1667	0.3448	0.5211	0.264	0.0848	0.0377	0.3074
4	0	0	0.1972	0.48	0.3152	0.1226	0.2278
5	0	0	0.0704	0.176	0.4	0.3019	0.1484
6	0	0	0.0141	0.04	0.1939	0.533	0.0904

93\92	1	2	3	4	5	6	Ergodic dist
1	0.3333	0.1538	0.0194	0	0	0	0.0453
2	0.6667	0.4872	0.2136	0.0263	0.0064	0	0.168
3	0	0.2564	0.4757	0.2105	0.0764	0.0265	0.2233
4	0	0.1026	0.2136	0.5	0.2994	0.1258	0.2695
5	0	0	0.068	0.2368	0.414	0.3046	0.1892
6	0	0	0.0097	0.0263	0.2038	0.543	0.1047

96\95	1	2	3	4	5	6	Ergodic dist
1	0.5294	0.125	0.0053	0	0	0	0.0667
2	0.4412	0.5703	0.1958	0.0385	0	0.027	0.2376
3	0.0294	0.2578	0.5556	0.2885	0.125	0	0.3194
4	0	0.0469	0.2116	0.4744	0.375	0.1081	0.2301
5	0	0	0.0317	0.1667	0.4063	0.2162	0.0989
6	0	0	0	0.0321	0.0938	0.6486	0.0474

97\96	1	2	3	4	5	6	Ergodic dist
1	0.6	0.1288	0.0208	0.0068	0	0	0.0827
2	0.3714	0.5379	0.1458	0.0473	0.0152	0	0.1936
3	0.0286	0.2727	0.599	0.3041	0.0303	0.0286	0.319
4	0	0.053	0.2083	0.4797	0.3182	0.0286	0.2215
5	0	0.0076	0.026	0.1486	0.5606	0.1143	0.1149
6	0	0	0	0.0135	0.0758	0.8286	0.0683

98\97	1	2	3	4	5	6	Ergodic dist
1	0.6279	0.1417	0.0101	0	0	0	0.1768
2	0.3256	0.675	0.2864	0.0709	0	0	0.4441
3	0.0465	0.175	0.5829	0.4397	0.1014	0.0278	0.2854
4	0	0.0083	0.1156	0.3546	0.4928	0.0833	0.0709
5	0	0	0.005	0.1206	0.3623	0.2222	0.0175
6	0	0	0	0.0142	0.0435	0.6667	0.0053

99\98	1	2	3	4	5	6	Ergodic dist
1	0.6739	0.1049	0.0048	0.009	0	0	0.1049
2	0.2826	0.6235	0.244	0.018	0	0	0.2936
3	0.0435	0.2407	0.5311	0.3063	0.1569	0.0345	0.3162
4	0	0.0309	0.201	0.5225	0.4314	0.1724	0.2085
5	0	0	0.0144	0.1351	0.2745	0.2414	0.053
6	0	0	0.0048	0.009	0.1373	0.5517	0.0238

00\99	1	2	3	4	5	6	Ergodic dist
1	0.6	0.1317	0	0.0076	0.0256	0	0.0897
2	0.34	0.503	0.2051	0.0379	0	0	0.2551
3	0.06	0.3293	0.641	0.4091	0.0769	0.08	0.4403
4	0	0.0299	0.1385	0.4394	0.4359	0.04	0.1569
5	0	0.006	0.0154	0.1061	0.3077	0.32	0.043
6	0	0	0	0	0.1538	0.56	0.015

01\00	1	2	3	4	5	6	Ergodic dist
1	0.6296	0.1849	0.0289	0	0	0	0.2438
2	0.3519	0.6438	0.3058	0.0833	0.0526	0	0.4526
3	0.0185	0.1712	0.5041	0.4815	0.1316	0.05	0.2279
4	0	0	0.1446	0.3796	0.3684	0	0.0607
5	0	0	0.0165	0.0556	0.3947	0.25	0.0127
6	0	0	0	0	0.0526	0.7	0.0022

02\01	1	2	3	4	5	6	Ergodic dist
1	0.7059	0.1162	0	0	0	0	0.155
2	0.2206	0.6162	0.3495	0.0556	0	0	0.3924
3	0.0735	0.2475	0.4903	0.4667	0.1333	0	0.316
4	0	0.0202	0.1553	0.4	0.3667	0.0625	0.1074
5	0	0	0	0.0778	0.4	0.25	0.0184
6	0	0	0.0049	0	0.1	0.6875	0.0108

Table 4: Markov transition matrices for 1990-2001 Ohio