

Notes on Poverty Traps and Appalachia

Steven N. Durlauf
University of Wisconsin at Madison

Preliminary Draft

September 25, 2009

In these notes, I provide some general ideas on how to conceptualize poverty traps and speculate on their applicability to understanding Appalachian poverty. My goal is to stimulate thinking on Appalachia that exploits contemporary perspectives in economics on the sources of persistent poverty and inequality. To do this, I focus on both the theory of poverty traps as well as issues in the econometric assessment of their empirical salience.

My discussion reflects the large body of modern literature on persistent poverty. Azariadis and Stachurski (2005) is an extensive survey of poverty traps with an orientation toward cross-country differences. With respect to national economies, poverty trap analysis has been important in understanding the continuing levels deprivation in much of the world. Durlauf (2004) provides an analogous survey from the vantage point of neighborhood effects and their attendant effects on inequality¹; ghettos are often regarded as a canonical example of a neighborhood level-poverty trap. The generative mechanisms for poverty traps of course differ according to the scale of the units under study. I conjecture Appalachia *may* be an example of a poverty trap, although regions per se are not the usual scale at which poverty traps are studied in modern research, there are aspects of poverty traps for both aggregate economies and for local neighborhoods that provide insights into regions.

While I am not aware of any modern research in economics that explicitly studies Appalachia from the perspective of poverty traps, ideas closely related to poverty traps have long been associated with the region. Over 30 years ago, Billings (1974)² described standard thinking on Appalachia:

¹Jencks and Mayer (1990) is a survey of social science research that precedes the modern economics literature. See also Manski (2000) and Durlauf and Ioannides (2009) for surveys of social interactions models, which focus on how groups affect individuals.

²Billings argues cultural explanations are overstated, a position maintained in Billings and Blee (2000).

The culture of poverty is the most common theory in the literature on poverty and Appalachia alike...Culture of poverty explanations, when applied to Appalachia, take several forms...Emphasis is on the debilitating effects of an atavistic, frontier culture and the socialization of its people into backwardness. The subcultural claim is often buttressed by an assertion that, with the recent introduction of improved roads and mass media, Appalachia is experiencing its first contact with the outside. (315-316)

Eller (2008) further argues that culture of poverty arguments helped motivate war on poverty strategies for Appalachia. One use of these notes, I hope, is to provide quantitative versions of culture of poverty arguments which will both clarify theoretical thinking as well as provide insights into how one can assess their empirical salience.

One important feature of modern poverty theories, whether defined at the aggregate or local levels, is their emphasis on the interplay of a large range of causal factors in producing (or eliminating) a poverty trap. This richness comes at a price as the empirical evidence for any particular factor is consequently difficult to assess. On the other hand, this richness is important in developing poverty trap perspectives that respect the heterogeneity of individual and subregional outcomes within Appalachia.

Section 1 of these notes provides a description of ways to conceptualize poverty traps. My goal here is to describe some probabilistic models of income dynamics which can generate persistent poverty as equilibrium outcomes. Some of the mechanisms that give rise to this persistence are also discussed. Section 2 discusses identification issues that arise in statistical analyses of poverty traps. Any claim that Appalachia is in fact a poverty trap will ultimately require grappling with identification; I argue that the problem is not insuperable. Section 3 applies some of these ideas to Appalachia in what is admittedly a speculative fashion. Section 4 concludes.

1. Income dynamics and poverty traps

In this section, I outline some baseline models of income dynamics and assess their equilibrium properties from the vantage point of how one might conceptualize a poverty trap. The reason for proceeding this way is that there is no accepted formal definition of a poverty trap. Rather, the term encompasses, I believe, three logically distinct and mutually compatible qualitative claims about the nature of poverty:

i. poverty is highly persistent,

ii. poverty is not “self-correcting” in the sense that some of the mechanisms that generate poverty are such that poverty can be perpetuated indefinitely,

iii. poverty is perpetuated by certain aggregate or collective features of the socioeconomic environment in which individuals make decisions.

The absence of either a formal or a scalar definition of a poverty trap does not, in my view, make the term unhelpful.³ In contrast, in my view formal definitions of poverty traps are of interest to the extent they capture these qualitative claims. One important distinction between poverty trap concepts *i* and *ii* as opposed to *iii* is that first two concepts refers to properties of the income process whereas the third refers to the mechanisms underlying the income process. What sorts of aggregate or collective mechanisms fall under *iii*? One type may derive from the way agents interact via markets. Another type may derive from direct interrelationships between agents. Here it is important to differentiate between individual interdependences that are adjudicated via the price system and those that are not. The state of the coal industry affects poverty via the standard processes of equilibrium wage determination. This contrasts with the idea that a given region is associated with social norms that reduce the value attached to education. The latter types of influences are often known as social interactions or neighborhood effects; I use the term social interactions here.

For expositional purposes, I focus on intergenerational income dynamics. To do this, I consider a sequence of family dynasties *i*. Issues of intermarriage, fecundity, etc. are ignored so each generation of the family is a single individual who lives 2 periods. The pair of indices *i,t* denotes a single adult, so person *i,t* was born at $t-1$; adding more elaborate lifetime structure does not matter qualitatively. Adult income is denoted by $y_{i,t}$; while income is not meant to summarize the individual’s socioeconomic status, it will be the basis of measuring whether or not

³For example, the lack of a precise single definition for poverty traps does not inhibit evaluating whether a given case is a poverty trap. Following an example in Taylor (1998), the absence of a clear definition of money does not diminish the meaningfulness that a dollar bill is an example of money.

the person is in poverty; for simplicity poverty is defined by an income less than or equal to y^{POV} . The mathematical structures I describe do not depend on the intergenerational context and may be applied to behavior within one lifetime. On the other hand, it is only in a particular substantive context such as intergenerational mobility that one can discuss the mechanisms that may underlie the algebra.

One view of intergenerational income dynamics is family-specific in that parental income determines offspring income. The formal analysis of this type of model was pioneered by Becker and Tomes (1979) and Loury (1981); a deterministic version of this class of models produces the simple law of motion for family income

$$y_{i,t} = \varphi(y_{i,t-1}). \quad (1)$$

It is standard to assume that $\varphi(\cdot)$ is nondecreasing in y . As this model is assumed to apply to all member of the population, this equation is sufficient to describe the evolution of the complete cross-sectional distribution of income over time, it allows one to characterize poverty rates, degrees of inequality and other population-wide aggregates.

From the vantage point of an individual family dynasty, an immediate implication of this structure is that for every initial condition $y_{i,0}$, income will either converge to a limiting value, i.e. steady state values of income y such that $y = \varphi(y)$, or diverge to infinity. Ignoring the latter possibility (which only adds technical complications at this point), the long run properties of the income dynamics process are fully summarized by the steady states of eq. (1).

Multiple steady states allow one to provide one formalization of the concept of a poverty trap. To do this, consider the properties of differences in income between two families i and j , i.e. $y_{i,t} - y_{j,t}$. If eq. (1) is associated with a unique steady state, then it is immediate that regardless of the value of the difference in incomes today

$$\lim_{T \rightarrow \infty} y_{i,t+T} - y_{j,t+T} = 0. \quad (2)$$

In words, the uniqueness of a steady state implies that any contemporaneous inequality will disappear over time. In contrast, suppose that there exist multiple steady states. Further, assume that these steady states are locally stable, which means that if a family starts sufficiently near either value, it will converge to that value. Finally, designating one of these stable steady states as y^L and another as y^H , suppose that relative to the poverty threshold y^{POV}

$$y^L < y^{POV} < y^H. \quad (3)$$

In this case, there exist levels of poverty and nonpoverty that are fully self-perpetuating. Eq. (3) thus constitutes one formalization of the idea of a poverty trap: moving from specific income values to ranges of incomes, families whose incomes lie in the vicinity of y^L will remain poor forever while families whose incomes lie far enough away from y^L will not. This captures the qualitative poverty trap ideas *i* and *ii*.

Under what conditions can (3) arise for dynamics (1)? Algebraically, the existence of a poverty trap requires that there exist income levels $y_1 < y_2$, such that $\varphi'(y_2) > \varphi'(y_1)$. If $\varphi(\cdot)$ is everywhere differentiable this condition requires that $\varphi'(\cdot) > 1$ for some values of $y_{i,t}$; it can also hold if there is jump discontinuity in $\varphi(\cdot)$. Intuitively, in order for this form of a poverty trap to occur, it is necessary that it is possible along an equilibrium path for families with higher incomes to grow faster than families with lower incomes. And what applies across families must also apply within families, i.e. it must be the case that for a given family, income growth is increasing with respect to initial income for some income levels. Notice that there is no requirement that income growth is increasing in income at all income levels. Hence, over a cross-section, one can observe an average tendency for income growth to be negatively correlated with initial incomes even though a poverty trap is present.

Eq. (1) is a reduced form description of equilibrium behavior and so is a black box in the sense that it describes the equilibrium dynamics of income for a family but does not explicitly describe the mechanisms by which income of a parent affects an offspring. In other words, the function $\varphi(\cdot)$ is determined in equilibrium by the underlying decision problems of parents. One mechanism that provides outcomes consistent with eq. (1) involves human capital formation and is

the one studied by Becker and Tomes and Louri. The economic logic underlying in these models is straightforward: parents divide income between consumption and human capital investment in children; these human capital investments, in turn determine income when children become adults⁴. If the level of investment in children is a nondecreasing function of income and the marginal product of human capital is positive, family income will evolve according to eq. (1) with the sign restriction we have imposed. Delineating this type of structure is important as it indicates that one must be careful in talking about causes of poverty traps. As the human capital explanation shows, there is an interplay between the preferences of parents (which determines the relationship between income and human capital investment) and technology (which determines the transformation of human capital into income).

While eq. (1) is consistent with the Becker and Tomes (1979) and Louri (1981) frameworks, poverty traps do not arise in either of their analyses. The reason for this is that each placed assumptions on the production function mapping human capital to income that in essence, ensured that $\varphi'(\cdot) < 1$ everywhere. Alternative specifications can produce different properties for $\varphi(\cdot)$ and hence generate poverty traps while preserving the behavioral foundations of their models. One way a poverty trap can occur is if the production function exhibits a region of increasing returns to human capital formation. A second way to produce a poverty trap in the family dynasty context is via lumpiness in human capital investment. If transitions across human capital levels require fixed costs to be paid, then $\varphi(\cdot)$ can exhibit a jump as the poor do not make these investments whereas the nonpoor do; Azariadis and Drazen (1990) is a classic example of a poverty trap driven by jumps. It is important to be clear that this sort of explanation in no way “blames the poor.” When preferences are homogeneous, the investment decisions of the poor are identical to those the more affluent would make in the same position.

The example of fixed human capital investment costs raises an important issue in the economics of poverty traps, namely the question of whether they require some sort of market incompleteness to sustain them. One reason concerns the ability of adults to borrow. Depending on the returns to human capital investment, poor families might wish to borrow in order to invest in their children and break a poverty trap. One impediment to borrowing of this type was first

⁴This formulation is similar to economic growth models in which aggregate economies build up capital stock via savings and consumption decisions.

recognized by Lounry (1981): parents cannot borrow against the future earnings of their children. Other types of financial market imperfections have been studied, see for example Galor and Zeira (1993). This type of explanation implies that there can be efficient redistribution of educational resources in the sense that equalization leads to greater aggregate output.

This first conceptualization of a poverty trap is fragile in an important sense. Suppose that one allows for randomness in incomes via a variable $\varepsilon_{i,t}$ which summarizes labor market luck, ability, etc., and modifies the income process from eq. (1) to

$$y_{i,t} = \varphi(y_{i,t-1}, \varepsilon_{i,t}). \quad (4)$$

Questions about poverty dynamics, traps, etc. of course immediately become probabilistic in such a context. For example, long run differences between families are more naturally described by calculations of objects such as the expected gap between two families in the future given the gap in their contemporary incomes, i.e.

$$\lim_{T \rightarrow \infty} E(y_{i,t+T} - y_{j,t+T} | y_{i,t} - y_{j,t}). \quad (5)$$

If this expected value is 0, one has a condition that is analogous to (2) above. Similarly, one can calculate the probability that a poor family will stay poor for the arbitrarily distant future, i.e.

$$\lim_{T \rightarrow \infty} \Pr(y_{i,t+T} < y^{POV}, y_{i,t+T-1} < y^{POV}, \dots, y_{i,t+1} < y^{POV} | y_{i,t} < y^{POV}). \quad (6)$$

If (6) equals 1, this is the equivalent of eq. (3) for a random environment. On the other hand, it is possible for (6) to lie between 0 and 1, which provides a richer notion of a poverty trap, i.e. a situation where a family is in danger of being poor for the indefinite future.

Income dynamics as generated by eq. (4) unsurprisingly exhibit very different properties from those implied by eq. (1). More surprising, the introduction of even a small amount of randomness can affect the existence of a poverty trap, i.e. even if (3) holds for a world without randomness, (6) can equal 0. The reason for this fragility is that the random term $\varepsilon_{i,t}$ can act to

overcome the effects of $y_{i,t-1}$ on a given individual. Repeated draws of $\varepsilon_{i,t}$ across time can, in turn lead to realizations so that even, if a poverty trap exists without shocks, the realizations cause a family to escape the trap. A simple algebraic example can illustrate this. Suppose that income is either high or low, i.e. there are only two possible values y^L and y^H and that income dynamics obey the Markov chain

$$\Pr(y_{i,t} = y^L | y_{i,t-1} = y^L) = \Pr(y_{i,t} = y^H | y_{i,t-1} = y^H) = 1. \quad (7)$$

Clearly this is an example of a poverty trap in the sense of (3). On the other hand, suppose that income is stochastic and follows

$$\Pr(y_{i,t} = y^L | y_{i,t-1} = y^L) = \Pr(y_{i,t} = y^H | y_{i,t-1} = y^H) = 1 - \delta. \quad (8)$$

No matter how small δ is, one can show that each family dynasty will spend, on average, one half of the time in poverty; further, any rank order in incomes between dynasties at one point in time will reverse itself with probability 1. Hence, no family is trapped in poverty and it is additionally guaranteed that any income differential between two families at one point in time will be reversed in the future.

There is another perspective along which one can construct a sturdier definition of a poverty trap, namely the expected number of generations before a poor family transits out of poverty. The expected number of generations before for this transition is $\frac{1}{\delta}$. As δ approaches 0,

$\frac{1}{\delta}$ diverges to infinity so this statistic replicates the notion of permanent poverty when there is no stochastic element. Expected passage times, in my view, are the more natural object of interest for empirical studies; put differently, a poverty trap as defined by (4) is a limiting and in certain ways idealized case of persistent poverty. In contrast, a poverty trap defined as a condition in which there exist long expected passage times out of poverty better respects heterogeneity in the effects of poverty on individuals as $\varepsilon_{i,t}$ is nothing more than unobserved individual-specific heterogeneity.

Therefore, letting $M^{P,NP}(y_{i,t})$ denote the expected value of the first passage time out of poverty for a family with initial conditions $y_{i,t} < y^{POV}$, then one can define families in a poverty trap as those for which

$$M^{P,NP}(y_{i,t}) \geq K. \quad (9)$$

The expected passage time before escaping poverty is, in my view, a natural statistic of interest if the objective of the analyst is to understand persistence, i.e. feature i of poverty traps. Of course, K needs to be specified by the analyst, but that is not a defect of the measure since it is a judgment as to how much persistence should be designated a trap. $M^{P,NP}(y_{i,t})$ itself of course does not require subjective judgment of this type. It is worth noting that calculations of this type are not uncommon in mobility analyses which focus on Markov transition processes. Occupational mobility is a standard context.

A third way to think about poverty traps, one that also permits a smooth transition between nonstochastic and stochastic environments, is to employ the structure of eq. (4) to uncover how initial conditions affect long run income levels. Assuming that the shocks $\varepsilon_{i,t}$ are uncorrelated across time (correlation in the shocks is trivial to handle as one simply works with the innovations in the moving average representation of $\varepsilon_{i,t}$ instead of $\varepsilon_{i,t}$ itself) one can construct a new time series

$$\tilde{y}_{i,t+T} = \varphi(\tilde{y}_{i,t+T-1}, 0) \text{ given } \tilde{y}_{i,t} = y_{i,t}. \quad (10)$$

The variable $\tilde{y}_{i,t+T}$ represents the family income levels that would occur under the counterfactual that all shocks starting at time t equal 0. The properties of this time series reveal the extent to which current income inequality is or is not self-correcting as it studies income dynamics after unpredictable future events are purged; as before this process will have a well defined limit

$$\tilde{y}_i^{\text{lim}}(y_{i,t}) = \lim_{T \rightarrow \infty} \tilde{y}_{i,t+T} \text{ given } \tilde{y}_{i,t} = y_{i,t}. \quad (11)$$

This limit is expressed as a function of income at time t ; it is possible that this limit is independent of its value. Elsewhere I have argued that if the limit in eq. (11) does depend on $y_{i,t}$ so that long run behavior depends on initial conditions (in this case income), this property captures what economic historians have meant by path dependence. With respect to poverty traps, one can modify (3) to define a poverty trap as the existence of income levels such that

$$\tilde{y}_i^{\text{lim}}(y^L) < y^{\text{POV}} < \tilde{y}_i^{\text{lim}}(y^H). \quad (12)$$

This concept corresponds to idea *ii*, that poverty traps involve the absence of self-correction mechanisms to overcome current poverty.

ii. location-based models

The family- or individual-specific perspective on income dynamics renders the location of the trap irrelevant. In other words, there is nothing about the community or region which matters for the trap; if the members of the population were redistributed across different communities or regions, the prospects would be unaffected, assuming that eq. (1) is a complete description.

A second class of income dynamics, one in which location matters, may be trivially generated by including vectors of location-specific factors $c_{l,t-1}$ and $c_{l,t}$ in the income dynamics process

$$y_{i,t} = \varphi(y_{i,t-1}, c_{l,t-1}, c_{l,t}, \varepsilon_{i,t}). \quad (13)$$

The vectors $c_{l,t-1}$ and $c_{l,t}$ separately appear in order to capture location influences that occur during childhood versus adulthood.

It is evident that the presence of these location effects can generate persistent poverty. So long as there is sufficient heterogeneity in $c_{l,t-1}$ and $c_{l,t}$ and sufficient sensitivity of φ to these vectors, then it is obvious families in different locations can exhibit differences in income over the

long term in the sense of (5) and that a family can be stuck in poverty in the sense of (6). In this respect, eq. (13) is able to simultaneously capture poverty trap concepts *i*, *ii* and *iii*.

Some elements of $c_{l,t}$ may be exogenous, canonical examples include geography and weather. Other locational factors may not literally be exogenous, but may evolve sufficiently slowly so that they may be treated as exogenous over time horizons over which poverty dynamics are evaluated. One example of this type include cultural norms or political institutions. For purposes of analyzing income dynamics, the key feature of $c_{l,t}$ is that its behavior may be taken as given without rendering the analysis incomplete. Manski (1993) refers to such factors as contextual effects, borrowing terminology from sociology.

A different way to introduce locational effects involves introducing feedbacks from the behaviors of members of the location to the behaviors of each individual. These feedbacks can occur over time or occur contemporaneously. Focusing first on intertemporal feedbacks, let $y_{-i,l,t-1}$ denote the vector of income levels for families in location l other than i at time $t-1$; introduction of this additional factor generalizes the income process to

$$y_{i,t} = \varphi(y_{i,t-1}, y_{-i,l,t-1}, c_{l,t-1}, \varepsilon_{i,t}). \quad (14)$$

It is common to assume that the average income of others $\bar{y}_{-i,l,t-1}$ is a sufficient statistic, so that

$$y_{i,t} = \varphi(y_{i,t-1}, \bar{y}_{-i,l,t-1}, c_{l,t-1}, \varepsilon_{i,t}). \quad (15)$$

Formulation (15) is the basis of an important class of formal models of poverty traps; Loury (1977) is a remarkable early version of this model. In terms of underlying economics, this dependence can occur because of local finance of public schools; see Bénabou (1996a,b) and Durlauf (1996a) for examples of formal analyses of neighborhoods and the transmission of poverty. Another source of dependence may involve role model effects. A second reason for such influences is informational. If adolescents make schooling choices such as effort on the basis of future economic benefits, the assessment of these benefits may depend on the distributions of educational levels and incomes observed in a community. Stratification of communities according

to income will correspondingly mean that different locations produce different inferences about the value of education. See Streufert (2000) for a complete analysis of this type of locational effect. Yet a third explanation relates to recent work on identity (Akerlof and Kranton (2000,2002)). Suppose that one effect of educational choices by an individual concerns how he relates his own identity to that of others in his community. If the link between education and identity depends on the characteristics of parents, then it is possible that (15) is an approximation to the effects of identity on choice. To be concrete, in a community where few parents are well educated, high education can render an individual feeling alienated from those with whom he wants to share an identity. This argument has of been long standing importance in understanding racial inequality as a number of authors have argued that black educational attainment is hampered by the perception that academic success is a form of “acting white” (Fryer and Torelli (2005), Ogbu (2003)). Suggestive evidence also exists of low aspirations among Appalachian youth, e.g. Ali and Saunders (2008).

A final modification of the income dynamics equation involves the introduction of contemporaneous locational influences, i.e. one allows for individual incomes to be affected by the current incomes of others as well as the current characteristics of a location so that

$$y_{i,t} = \varphi\left(y_{i,t-1}, \bar{y}_{-i,l,t-1}, \bar{y}_{-i,l,t}, c_{l,t-1}, c_{l,t}, \varepsilon_{i,t}\right). \quad (16)$$

When choices are discrete, this model thus corresponds to social interactions models of the type studied by Brock and Durlauf (2001a,2006,2007) and other authors; see Durlauf and Ioannides (2009) for a recent survey. One source for contemporaneous income interdependences is informational: to the extent that labor market information flows across social networks, economic success by a member of one’s network can mean greater information. Empirical evidence of this phenomenon is developed by Topa (2001) and Bayer, Ross and Topa (2008); the first paper also provides a formal theoretical model of information transmission.

From a theoretical perspective, the introduction of $\bar{y}_{-i,l,t}$ is especially interesting as its presence means that the model has the capacity to produce multiple equilibria in the density of incomes within a given location. In terms of thinking about poverty traps this is of particular importance as it creates the possibility that two locations with identical distributions of individual and locational characteristics can exhibit different levels of aggregate income. How can this

happen? Suppose one considers work effort rather than income as the object of interest. If the productivity of effort is complementary in the effort of others, i.e. the marginal product of increased effort by one worker is increasing in the effort levels of others, then the effort choices of each individual will be increasing in the effort levels of others⁵. If this complementarity is strong enough, then there will exist multiple effort levels across a population; each of these distributions of effort levels is self-consistent in the sense that they represent Nash equilibria: no one has an incentive to change his effort level given the choices of others.

When can multiple equilibria occur? Brock and Durlauf (2001a,2006) show, for discrete choices, that the number of equilibria in a given economic context depends on the interplay of private and group characteristics with the strength of complementarities. Intuitively, if private incentives polarize the population toward one type of behavior, then complementarities cannot create sufficient bunching so that the population on average tips to the other choice. Further, if the distribution of individual-level unobservables generates large draws of the epsilons' with sufficient frequency, then the percentage of the population left over to react to other factors will be insufficient to generate multiple distributions of self-consistent bunching. This formalizes the idea that sufficient iconoclasts in a population can break socially-enforced conformity among others. As was the case for the individualistic models of poverty traps, the robustness of a poverty trap for a social model of this type depends on the ways shocks impinge on individual decisions. Unlike the individualistic income dynamics model (1), however, what matters in the locational model is whether enough aggregate heterogeneity is induced to overcome the potential of strong conformity effects tipping the rest of the population.

iii. growth

The discussion up to this point has focused on environments in which incomes do not systematically grow. The modern economic growth literature has focused on cases where interactions can lead to perpetual and endogenous growth. The basic idea, initially proposed in

⁵ Formally, for a function $f(r, s, \dots)$, r and s are complementary if $\frac{\partial f(r, s, \dots)}{\partial r \partial s} > 0$. In words, the marginal effect of increasing one variable is itself an increasing function of the level of the other. One can extend this definition to vectors and generalize to cases where functions are not twice-differentiable.

seminal work by Lucas (1988) and Romer (1986), is that the productivity of capital investments, whether human or physical, depends on the investments of others in the economy. Applying these ideas to the Appalachian case, one would say that the effect of human capital on one worker's productivity is increased by greater human capital on the part of others. Lucas and Romer sought to explain long run divergence between developed and lesser developed economies and so focused on the case where these spillovers produced "social increasing returns to scale" which means that if the capital levels of others is fixed, a given individual faces a decreasing returns to scale mapping of capital into income, while aggregate economy exhibits increasing returns. For our purposes, their model is a variant of (16) in which the individual income variables grow without bound.

In terms of the conceptions of poverty traps I have described, perpetual growth requires a modification of the various formalizations. One possibility is to think of traps in terms of relative versus absolute deprivation. This would involve considering the behavior of variables such as $\log y_{i,t} / y_{j,t}$; the use of logs accounts for the idea that in growing economies, a fixed difference in income becomes a negligible fraction of the incomes. A relative deprivation trap could then be conceptualized as one in which contemporaneous inequality can be permanent, i.e.

$$\Pr\left(\lim_{T \rightarrow \infty} \left(\log\left(y_{i,t+T} / y_{j,t+T}\right) \geq K, \log\left(y_{i,t+T-1} / y_{j,t+T-1}\right) \geq K, \dots, \log\left(y_{i,t+1} / y_{j,t+1}\right) \geq K \mid \log\left(y_{i,t} / y_{j,t}\right) \geq K \right)\right) \quad (17)$$

is positive. The factor K of course needs to be set just as the definition of absolute poverty had to be set in our earlier definitions of poverty traps in absence of growth.

One question is whether the sorts of generative mechanism that produce social increasing returns to scale in aggregate economies apply to regional economies. Romer and Lucas put much emphasis on idea generation. Lucas (2009) argues

What is it about modern capitalist economies that allows them, in contrast to all earlier societies, to generate sustained growth in productivity and living standards?...What is central, I believe, is the fact that the industrial revolution involved the emergence (or rapid expansion) of a *class* of educated people, thousands, now many millions-of people who spend entire careers exchanging ideas, solving work-related problems, generating new knowledge. (1)

One can see analogies to regional development in this statement. Bollinger, Ziliak, and Troske (2009) find evidence that returns to education are lower in Appalachia than the rest of the United States, which is consistent Romer-Lucas type spillovers, although this fact would arise wherever education levels of workers are complementary in production functions.

iv. A summary statistical model

In using any of these dynamic income models to assess data, it is evident that there are forms of heterogeneity that are missing. One would expect that there exist a range of individual specific variables that affect outcomes, denote these as $x_{i,t}$; as before one may choose to distinguish between influence in youth versus adulthood. In empirical work, these represent individual-specific observables. Finally, it is necessary to allow for locational specific unobservables in both youth and adulthood. If we define the location specific unobservables as $\eta_{l,t}$, one has a general process for individual income of the form

$$y_{i,t} = \varphi(y_{i,t-1}, x_{i,t-1}, x_{i,t}, \bar{y}_{-i,t-1}, \bar{y}_{-i,t}, c_{l,t}, c_{l,t-1}, \varepsilon_{i,t}, \eta_{l,t-1}, \eta_{l,t}). \quad (18)$$

This general specification respects the distinctions between individual and locational influences, observable and unobservable heterogeneity, and contextual and endogenous factors. As such, it naturally corresponds to the type of statistical model one would apply to individual income dynamics.

Along some dimensions, this model is easily generalized. Eq (18) can immediately be extended to a vector of individual outcomes, so that the coevolution of other socioeconomic outcomes such as human capital are objects of interest in addition to income. As the above discussion indicates, it is often the case that one thinks of various social interactions operating with respect to outcomes other than income per se. There is another dimension, however, with respect to which this formulation is incomplete: it says nothing about why individuals live in particular locations. In the case of country-wide poverty traps, this is not an important lacuna given international immigration restrictions, but in the case of the United States, this needs to be considered. In the poverty trap literature, the standard explanation as to why poorer families do

not move to locations that will maximize human capital in their offspring is that housing prices and rents sustain substantial socioeconomic segregation. When one considers racial inequality, discrimination may act as a separate barrier; see Yinger (1995) for evidence on housing discrimination and Heckman (1998) for a critique of this work.

Before turning to econometric issues, it is worth observing that location-specific factors raise question of interventions to affect the allocation of individuals across localities. In Durlauf (1996b) I have termed this associational redistribution. Many locational factors act as externalities in the sense that they are not directly adjudicated by markets; peer group effects are a standard example. This is so even if prices (i.e. house prices or rents) for locations support the allocation of agents; see Becker and Murphy (2000) for a very clear treatment and Bénabou (1996a) for detailed analysis. Hence, it would seem that there can be efficient interventions in market allocations of individuals across locations. On the other hand, the presence of complementarities between characteristics of agents can render stratification by these characteristics efficient; this is Becker's classic (1973) result on the efficiency of assortative matching, i.e. stratification of groupings. One can identify cases where complementarity does not render assortative matching efficient (Prat (2002), Durlauf and Seshadri (2003)); Bénabou (1996b) is a standard reference for studying the efficiency of stratification in the context of school districts when complementarities occur at both local and aggregate levels. Nevertheless, Becker's basic message delimits the probable efficiency gains from government interventions in group formation that are designed to equalize agent characteristics across groups. Location-driven poverty traps may therefore represent an example in which one may have to trade off equality against efficiency.

2. Identification

In this section, I discuss the question of identifying poverty traps. The objective of the discussion is to both illustrate the identification problems that arise in producing elements of poverty traps as well as to describe strategies for overcoming these problems. To make the general econometric issues concrete, I consider a specific empirical proposition and its interpretation. Suppose one argues that Appalachia's historically high poverty rates as compared to the rest of the country represent *prima facie* evidence of a poverty trap. I would argue that the empirical

regularity, i.e. the fact of historically high poverty rates, does not necessarily constitute evidence of a poverty trap with respect to *any* of the ideas underlying the poverty trap notion. Idea *i*, persistence in individual level poverty is not demonstrated by the empirical regularity for an obvious reason: persistence in individual poverty does not logically restrict aggregate poverty levels in a location. This follows immediately from the fact that the percentage of a population in poverty does not identify anything about the dynamics of individual poverty processes⁶. Further, one can think of a plethora of reasons why high aggregate poverty rates in a region would be uninteresting from a policy perspective. One reason is migration; persistent poverty in a location can reflect location decisions of agents who would be poor regardless of location; by analogy, the concentration of poor in low quality housing does not imply that low quality housing is a poverty trap. Idea *ii*, absence of self-correction of poverty cannot be deduced from high aggregate poverty for exactly the same reasons. Idea *iii*, the presence of aggregate reasons for individual poverty, does not follow from persistent poverty either. Here the reason is simple: nothing in the aggregate poverty fact speaks to its causes. One cannot tell from high aggregate poverty whether it is due to low family specific investments in human capital because of individual family poverty, a weak tax base for public education, absence of incentives to invest due to the state of the coal industry, particular social norms about education or other factors. Taking poverty trap ideas seriously requires much more detailed knowledge about individual income dynamics.

i. evaluating poverty traps via time series properties

One strategy for generating evidence of poverty traps may be derived from explicit consideration of the time series properties of individual income dynamics. Calculations of this type directly address the phenomena of persistence and absence of self-correction in poverty. As suggested above, it is important to allow for nonlinearities in the transition function. Despite their commonality in theoretical work there has been relatively little empirical work on the question of nonlinearities in the intergenerational transmission mechanism. For the United States, exceptions include Cooper, Durlauf and Johnson (1994), who find little correlation between parental and offspring income outside the tails of the income distribution. A particularly careful analysis for

⁶The mathematical point is that poverty in a region is a description of the cross-section density of incomes at a point in time, which does not map one for one back to a particular dynamic process for the incomes.

non-US data is Antman and McKenzie (2007) which estimates nonlinear intergenerational models for 15 years of individual data on urban Mexican workers and finds little evidence of nonlinearity.

In translating the statistical notions of poverty traps into econometric analysis, it is important to recognize an important limitation: namely, while poverty traps require the possibility that the rich grow faster than the poor, the observation that the poor in fact grow faster than the rich does not imply the absence of a poverty trap. The reason can be seen in the nonstochastic version of the individual-based poverty trap model described by eqs. (1) and (3): in the vicinity of each of the steady states described in eq. (3), one has the property of local convergence, i.e those below the steady state grow towards it while those above it shrink. Bernard and Durlauf (1996) discuss this problem. One implication of their analysis is that linear models of income dynamics cannot be used to assess poverty persistence; specifically it is possible to find a cross-sectional correlation between initial income and income growth in an environment with poverty traps because this correlation does not account for any nonlinearities associated with the poverty trap.

A second identification problem concerns nonlinearity versus poverty traps. Following Durlauf, Johnson, and Temple (2005), the difficulty arises in the relationship between clustering of behaviors around a discrete number of values and the presence or absence of multiple steady states. Clustering implies that there are few observations that are not associated with the clusters and hence uncovering transition dynamics towards clusters is difficult. In the context of our model, the problem can occur because of a lack of information about behaviors around the discontinuities in (1); for the continuous case the problem would arise if the set of income over which $\varphi'(\cdot) > 1$ is small. This reinforces the importance, in my view, of focusing on transition times out of poverty; shorter transition times are presumably reasonably easy to estimate.

A distinct problem arises with respect to the accuracy of estimates if one treats the definition of a poverty trap as requiring permanent poverty. Such a stark requirement is difficult to assess from data which are observable over a relative small epoch, say 50 years. In the time series literature, this problem has arisen in the context of the study of unit roots in macroeconomic data. A unit root in a time series requires that some part of the contemporary change in a time series permanently affects the level of a time series, so there is a close relationship to the poverty trap claim that a change can leave someone in or out of a trap. For income y_t , the expected long run implication of a contemporaneous change in income may be calculated via

$$\lim_{T \rightarrow \infty} E(y_{t+T} | \Delta y_t) = \sum_{j=-\infty}^{\infty} \sigma_{\Delta y}(j) \quad (19)$$

where $\sigma_{\Delta y}(j) = \text{cov}(\Delta y_t, \Delta y_{t-j})$. Hence calculating permanent effects involves high order covariances, which are extremely difficult to estimate accurately with extremely long samples, an issue first assessed in Cochrane (1988). For the unit root context, this means that permanence cannot be assessed nonparametrically with any degree of accuracy. So, if one formalized the notion of a poverty trap as requiring that some transformation of aggregate poverty rates exhibits a unit root, evidentiary support will be problematic. The same holds for other conceptions of poverty traps; the semiparametric analyses such as Cooper, Johnson, and Durlauf and Antman and McKenzie avoid this problem by focusing on transitions across a single generation which rules out any higher order temporal dynamics. In my view, this problem reinforces the importance of focusing on aspects of passage times out of poverty.

ii. locational mechanisms

A different strategy for uncovering poverty traps is to focus not on time series regularities, but on the identification of feedbacks from various locational characteristics that correspond to contextual and endogenous social interaction influences on individual outcomes. From this viewpoint, the objects of interest are the derivatives of eq. (18) with respect to the social interactions variables. If one can uncover these derivatives, one can infer poverty trap outcomes in the sense of *iii*. This is the strategy that is employed in the social interactions literature. The most common social interactions models are linear regression variants of (18); if one were to map (18) into a linear regression it would take the form

$$y_{i,t} = \kappa + \alpha_1 x_{i,t-1} + \alpha_2 x_{i,t} + \beta_1 c_{l,t-1} + \beta_2 c_{l,t} + \gamma_1 \bar{y}_{-i,l,t-1} + \gamma_2 \bar{y}_{-i,l,t} + \eta_{l,t-1} + \eta_{l,t} + \varepsilon_{i,t}. \quad (20)$$

The most important variant of this model is one in which choices are discrete variables. For simplicity, I focus on the binary choice case; denote these outcomes $\bar{y}_{i,t} \in \{0,1\}$. In this approach, the net utility to choice 1 by agent i at t , $u_{i,t}$, obeys an analog to (20)

$$u_{i,t} = \kappa + \alpha_1 x_{i,t-1} + \alpha_2 x_{i,t} + \beta_1 c_{l,t-1} + \beta_2 c_{l,t} + \gamma_1 \bar{y}_{-i,l,t-1} + \gamma_2 \bar{y}_{-i,l,t} + \eta_{l,t-1} + \eta_{l,t} + \varepsilon_{i,t} \quad (21)$$

so that the observed behavior follows

$$y_{i,t} = 1 \text{ if } u_{i,t} > 0; y_{i,t} = 0 \text{ otherwise.} \quad (22)$$

My specifications of both the linear regression and binary choice models are more complicated than the statistical models that have usually been employed to study social interactions. Note that the specifications assume that the elements of x and c are known. This can be problematic since theoretical models of individual and location determinants typically fail to specify how determinants should be measured. For role models, is the correct variable the percentage of white collar jobs among adults or the percentage of college graduates among adults? This sort of question can be repeatedly asked. Further, the specifications follow the literature in taking locations that define social interactions as known. As argued in Akerlof (1997), it is natural to think about agents arrayed in a possibly high dimensional social space; this may or may not correspond well to counties, neighborhoods, which define the locations over which measurement is conventionally done.

The econometrics literature distinguishes between three distinct identification problems in uncovering locational influences from statistical models of the type I have described. Durlauf and Ioannides (2009) provide formalizations of the problems as well as an exhaustive description of the literature. Here, I simply wish to describe the problems a researcher faces.

The first identification problem facing studies of social influences was first studied in Manski (1993) and is known as the reflection problem. The reflection problem refers to difficulties in disentangling the role of contextual effects $c_{l,t-1}$ and $c_{l,t}$ from the endogenous effects $\bar{y}_{-i,l,t-1}$ and $\bar{y}_{-i,l,t}$. This difficulty arises because the contextual effects help to determine equilibrium values of the endogenous effects. Manski provides a demonstration that for cross-section linear models, the reflection problem may render it impossible to identify different locational effect parameters. Brock and Durlauf (2001a) show that the reflection problem does not arise in discrete choice models in the sense that because these models are nonlinear,

collinearity between contextual and endogenous effects may be ruled out, so long as there is sufficient variability in the contextual effects across locations. Brock and Durlauf (2001b) shows that the reflection problem can also be overcome in dynamic contexts because dynamics can affect the degree of linear dependence between the contextual and endogenous effects. Nevertheless, even if identification does not fail per se, the reflection problem indicates that parameter estimates may be highly imprecise.

A second econometric problem derives from self-selection into locations. In terms of the underlying econometrics, self-selection means that

$$E\left(\varepsilon_{i,t} \mid x_{i,t-1}, x_{i,t}, c_{l,t-1}, c_{l,t}, \bar{y}_{-i,t,t-1}, \bar{y}_{-i,t,t}, \eta_{l,t-1}, \eta_{l,t}\right) \neq 0. \quad (23)$$

The economic reasoning underlying these conditional expectations has been standard since Heckman's pioneering work: if agents choose locations, then knowledge of the chosen location will provide information about an individual's unobserved heterogeneity, since that heterogeneity will interact with other factors producing the choice.

Following the broader microeconometrics literature, self-selection in locations is typically addressed in two ways. First, instrumental variables may be employed. An early and well known application of this strategy is Evans, Oates and Schwab (1992) who studied social interactions in schools. In order to address self-selection in schools, Evans Oates and Schwab used school district level instruments, arguing that self-selection is limited to schools within districts, and not the districts per se. This example reveals some of the difficulties in using instrumental variables. Even if the Evans, Oates and Schwab argument on self-selection is correct, this is not sufficient to ensure instrument validity. The problem is that $\varepsilon_{i,t}$ contains all factors that are not accounted for by the locational and individual-specific controls. In order for a district-level instrument to be valid, one must be able to argue that it is not correlated with any of these factors. As a mathematical statement, the presence of $\eta_{l,t}$ is sufficient to make this impossible, except for nongeneric cases. Substantively, the problem is what Brock and Durlauf (2001c) have called theory-ependendness: models such as eqs. (20) and (21) are not derived from full specifications of individual decision problems and therefore do not rule out determinants outside of those that are included. For the Evans, Oates and Schwab context, it is not unreasonable to conjecture that an instrument such as

district level dropout rates is correlated with per pupil expenditure or broader social norms that affect decisions.

A second strategy is to explicitly model the self-selection process. In turn, there are two ways to proceed. One approach involves coupling the outcome equations (20) and (22) with models of the locational selection process; Epple and Sieg (1999) is a nice example. The other approach, one which requires less a priori information on the location choice process, is due to Heckman (1979) and involves introducing regressors, which are estimates of the conditional expectation in (23) modulo a constant of proportionality. These regressors control for (23) and this is in fact known as the control function approach. While implementation of the control function approach is most often done using parametric assumptions on the probability density of unobserved heterogeneity, there are semiparametric ways to construct selection corrections.

Either variant of this second strategy is, in my view, preferable to the use of instrumental variables. In particular, explicit analysis of self-selection can assist in the identification of social interactions. Brock and Durlauf (2001b) first demonstrated that it was possible for the reflection problem to preclude identification when individuals are randomly assigned to locations while if locations were chosen, identification was possible; Brock and Durlauf (2006) and Ioannides and Zabel (2008) extend this approach theoretically with Ioannides and Zabel applying it successfully to demonstrate the presence of social interactions in housing valuation. Why would self-selection facilitate identification? Selection of locations constitutes an additional choice on the part of individuals, and so contains information on the determinants of these choices, determinants that presumably include the social interactions that will be experienced conditional on residing in the location. This information can help to triangulate the presence of social interactions to the extent that the interactions influence his locational choices.

A third identification problem derives from the presence of unobserved location-level heterogeneity, i.e. the presence of $\eta_{l,t-1}$ and $\eta_{l,t}$ in (20) and (21). In my judgment, the identification of social interactions effects in the presence of unobserved group effects represents the major existing impediment to developing evidence of the role of social influences. First, it is generally the case that for those contexts in which social interactions are usually studied, there are many unobserved group characteristics that can be plausibly argued to affect individual outcomes. For Appalachia, factors ranging from the quality of legal and political institutions to geography plausibly matter in explaining poverty, but are difficult to measure. Second, unlike the case of self-

selection, unobserved group factors do not themselves typically derive from a behavioral model the way that location selection does. Hence, there is nothing analogous to the control function approach that may be employed to address their presence. Most efforts to address unobserved group effects have therefore involved instrumental variables methods or, when the effects are time invariant, differencing of data to exploit temporal variation.

For the reasons I have outlined, there continues to be considerable disagreement about the empirical importance of social interactions. Recent econometric work has focused on uncovering robust evidence. By robust evidence, I mean evidence of social interactions that explicitly accounts for the presence of various types of unobserved individual and locational heterogeneity.

One approach to developing robust evidence is due to Brock and Durlauf (2007) for binary outcomes and in essence does the following. Suppose that one observes that there exist two locations, l and l' and a vector z which is a sufficient statistic for the effects of individual and contextual characteristics on the aggregate locational outcome. What I mean by this is that the only factors that determine the average choice levels outside of z are unobserved group effects and endogenous social interactions. Suppose that one observes

$$\bar{y}_{l,t} > \bar{y}_{l',t} \text{ and } z_{l,t} < z_{l',t}. \quad (24)$$

Brock and Durlauf call this a pattern reversal: the basic idea is that the observable fundamentals suggest one rank ordering of locational outcomes whereas the observed pattern of outcomes reverses this ordering. Under the behavioral model I have described, a pattern reversal can occur because of 1) the group effects $\eta_{l,t}$ reverse the rank order in outcomes generated by $z_{l,t}$ or because there are multiple equilibria in aggregate outcomes, so that l has coordinated on a high outcome equilibrium whereas l' has coordinated on a low outcome equilibrium. When can the first explanation be ruled out? Brock and Durlauf provide a set of shape restrictions on $\eta_{l,t}$ such that explanation 2 must be the reason. For example, if higher z locations draw from a more favorable η distribution, then an observation consistent with eq. (24) allows one to conclude that endogenous social interactions are present in the data and that they are strong enough to produce multiple equilibria. As such, this is a form of a partial identification argument.

The appeal of the pattern reversal approach is that it focuses on the one feature of endogenous social interactions that other factors simply cannot produce: multiple equilibria. Other shape restrictions can also be used. One example is the following. Suppose that the unobserved heterogeneity is drawn from a unimodal density. Suppose that one finds that there exists a vector π such that conditional on \bar{y} , πz is multimodal. In other words, multiple equilibria are implied if, across locations with common outcomes, the fundamentals associated with them are bimodally distributed. In this case, one can again conclude that social interactions are present and strong enough to produce multiple equilibria. The disparate z 's that are associated with the same aggregate outcomes reflect the different equilibria that can occur for a given group.

A second strategy is due to Graham's (2008) extension and generalization of work by Glaeser, Sacerdote, and Scheinkman (1996) on the impact of endogenous social interactions on the variance of average outcomes across locations. In this work, one considers the relationship between the variance of $\bar{y}_{l,t}$ and the population size of l . If there are no endogenous social interactions, then this relationship will be different than when endogenous social interactions are present. Intuitively, endogenous social interactions introduce dependencies across individual choices that "slow down" the rate at which the law of large numbers applies. Glaeser, Sacerdote and Scheinkman's analysis does not allow for group effects; Graham's achievement is to show that if these effects are random rather than fixed, and if the variance is independent of group size, one can uncover evidence of endogenous effects by contrasting variances across group sizes. The random effects assumption implicitly requires that location choices are unaffected by their presence.

A third strategy is proposed in Brock and Durlauf (2009) and involves studying transitional dynamics. Their framework considers adoption of a technology, for an educational context one can think of the development of skills to use a new technology, for example computers. Brock and Durlauf ask what sorts of restrictions are imposed on adoption over time, when the benefits to a technology reflect social interactions. For their model, they show that social interactions can introduces jumps in adoption rates in the economy as well as pattern reversals between adoption rates of those who private characteristics would suggest they should adopt earlier versus others. To be concrete, suppose that one observed that computer technology diffused more slowly in a high education location versus another but that there are jumps in the adoption rates of each

community considered in isolation. Brock and Durlauf in essence provide sufficient conditions under which one can conclude social interactions are present.

To be clear, none of these strategies is a panacea. Each requires substantive behavior assumptions. Hence their utility can only be assessed in a specific empirical context.

iii. data

My discussion of identification has focused on statistical tools as opposed to data collection. One can well imagine that the evidentiary support for social interactions in one sphere can be identified from their presence in others. I conjecture that language use is one direction to pursue. It is well understood that dialects are important sources of identity, cf. Wolfram and Schilling-Estes (2006). Nonstandard dialects in Appalachia have been a major topic in the sociolinguistics literature; Wolfram and Christian (1976) is an especially detailed study. I believe that language similarity may help provide insights into the appropriate metric for measuring closeness in social space and may further represent a marker that helps determine to what extent identity is locally driven. Luhman (1990) is an interesting study that considers how standard dialect speakers in Kentucky form stereotypes about nonstandard dialect speakers in Kentucky as well as the extent to which these stereotypes are accepted.

3. Footprints of poverty traps

In this section I focus on some stylized facts about Appalachia that would seem to hint at poverty traps. While this discussion is admittedly speculative, it reflects impressions I have gleaned from studies of Appalachia as to ways to uncover social interactions that are strong enough to produce poverty traps under the various conceptions that have been described.

i. education

The stylized facts on Appalachian educational attainment are suggestive, in my view of a poverty trap. Isserman (1996) is one example of a literature that documents how Appalachia

appears to be an outlier in terms of the high percentage of counties in which less than $\frac{1}{2}$ the 1990 adult population graduate from high school. Bollinger, Ziliak and Troske (2009) similarly attribute much of Appalachian wages to converge to those in the rest of the country to lower human capital formation; this paper is noteworthy for its careful and sophisticated econometrics. Low human capital investment in Appalachia is a good candidate for a mechanism underlying a poverty trap. As discussed earlier, the sort of behavioral explanation that one can make is that educational investment decisions are interdependent because of both role model and peer influences, so that factors such as parental education and the educational choices of peers affect each individual's decisions. Evidence of social interactions in education have been developed in many studies. Crane (1991) is an early example in which interneighborhood variations in high school graduation are associated with the occupational characteristics of parents. Recent examples include Cooley (2008) and Hanushek, Kain, Markman and Rivkin, (2003) who focus on peer effects at the school and classroom level. For these reasons, I regard the fact of sustained disparities in education to be a hint of a poverty trap. Nevertheless, by itself, the social interactions/poverty trap interpretation falls prey to the sorts of identification problems I have described.

To proceed, consider two other stylized facts. The first is identified in Isserman: for socioeconomic indicators other than education, it is much more difficult to identify Appalachia as an outlier relative to the rest of the country. The finding that high school completion behaves differently from other socioeconomic indicators is potentially of great importance in uncovering why it occurs. A second stylized fact is due to Shaw, deYoung, and Rademacher (2004) who find that the bulk of the Appalachian educational gap is due to central Appalachia. This is most starkly seen in terms of high school graduation: in 2000 76.8% of Appalachian adults had high school degrees as opposed to 80.4% for the US as a whole. In contrast, only 64.1% of central Appalachian adults are high school graduates. The high dispersion of education outcomes in Appalachia across subregions provides the sort of variability that helps uncover social interactions.

In what sense might these additional facts help produce evidence of poverty traps? With respect to Isserman, the anomalous behavior of education versus other socioeconomic indicators makes an explanation based on unobserved location factors less plausible. The reason is simple: the unobserved factor will need to be one that only affects education, since it evidently does not affect other factors. While this may apply to teacher quality, it does seem plausible from the perspective of social norms. As for Shaw, deYoung, and Rademacher (2004), if it is the case that,

assessing county by county, one finds that the low educational attainment associated with central Appalachia violates patterns of education as would be predicted by variables I have described by $z_{l,t}$ above, this would constitute a pattern reversal. Interpretation of these reversals as social interactions would require taking a stance on unobserved group heterogeneity. If the relevant factor is teacher quality, it is plausible to assume that teacher quality is drawn from a distribution that is no better for high outcome counties than others. Alternatively, one might wish to assume unimodality of the unobservables and see if one finds conditional multimodality in outcomes. The Graham approach can also be used if one can argue that the unobservables are uncorrelated with z . Lichter and Campbell (2005) document sufficient heterogeneity in poverty reductions in the 1990's to suggest this route may be informative.

ii. migration

Second, I conjecture that substantial information on social interactions can be gleaned from understanding the determinants of migration in and out of Appalachia. Whether or not Appalachia is a poverty trap, socioeconomic conditions would lead one to expect substantial migration away from the region. Actual migration patterns are in fact much more complicated. Obermiller and Howe (2005) document that in the latter 1990's Appalachia experienced substantial inflows and outflows of population. Underlying these flows are important differences between in and out migration. Obermiller and Howe find that central Appalachia experienced nontrivial outflows of more skilled adults that were largely counterbalanced by inflows of lesser skilled ones; more generally Baumann and Reagan (2005) argue that slightly over 1/8 of the gap in college graduates between Appalachia and the rest of the United States can be attributed to migration.

There appear to be puzzles in the migration patterns that warrant study in terms of what they say about social interactions. One puzzle, at least to me, is that out migration of the higher skilled has not been more rapid, especially in light of findings such as Bollinger, Ziliak and Troske (2009) that returns to human capital are lower in Appalachia than elsewhere. Nor is it clear why lesser skilled workers would choose Appalachia as a destination. The retention of higher skilled workers suggests the presence of social interactions effects that make Appalachia more appealing than its observed socioeconomic characteristics would suggest. On the other hand, the in

migration of lesser skilled workers suggests that self-selection issues exist with respect to the Appalachian population that mitigate against claims of the region being a poverty trap per se. My point is that analysis of migration decisions can augment social interactions analyses based on outcomes data for Appalachian residences. Again, as documented by Obermiller and Howe, there are interesting migration patterns across Appalachian counties which are related to their economic status. Comparisons with migration patterns for other disadvantaged regions may also be informative.

4. Conclusions

In these notes, I have tried to do three things. First, I have described some of the income dynamics models that can produce behaviors that capture various facets of the idea of poverty traps. Second, I have discussed some of the statistical challenges facing any effort to establish the presence of a poverty trap in a given data set. Third, I have used Appalachian educational attainment as an example of where one might wish to begin a systematic search for evidence in light of some established aggregate regularities.

I end these notes with a few comments on policy. Poverty trap theories and the associated econometrics are sufficiently primitive that to use either to make strong policy recommendations is inappropriate. At best, each represents a way of buttressing the claim that a poverty trap is present. This means that qualitative work can play a useful and complementary role in assessing evidence on poverty traps. It also means that policy construction should place strong emphasis on robustness, i.e. policies should respect the ignorance of the policymaker. This does not lead to Hayekian avoidance of policy interventions, but rather leads to the need to construct policies that respect the degree of policymaker ignorance.

Bibliography

- Akerlof, G., (1997), "Social Distance and Social Decisions," *Econometrica*, 65, 5, 1005-1027.
- Akerlof, G. and R. Kranton, (2000), "Economics and Identity," *Quarterly Journal of Economics*, 115, 3, 715-753.
- Akerlof, G. and R. Kranton, (2002), "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature*, 40, 4, 1167-1201
- Ali, S. and J. Saunders, (2008), "The Career Aspirations of Rural Appalachian High School Students," *Journal of Career Assessment*, 17, 2, 172-188.
- Antman, F. and D. McKenzie, (2007), "Poverty Traps and Nonlinear Income Dynamics with Measurement Error and Individual Heterogeneity," *Journal of Development Studies*, 43, 6, 1057-1083.
- Azariadis, C. and A. Drazen (1990), "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 105, 2, 501-526.
- Azariadis, C. and J. Stachurski, (2005), "Poverty Traps," in *Handbook of Economic Growth*, P. Aghion and S. Durlauf, eds., Amsterdam: North Holland.
- Baumann, R. and P. Reagan, (2005), "The Appalachian Brain Drain," mimeo.
- Bayer, P., S. Ross and G. Topa, (2008), "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 116, 1150-1196.
- Becker, G., (1973), "A Theory of Marriage, Part I," *Journal of Political Economy*, 81, 4, 813-846.
- Becker, G. and K. Murphy, (2000), *Social Economics*, Cambridge: Harvard University Press.
- Becker, G. and N. Tomes, (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 6, 1153-1189.
- Bénabou, R., (1996a), "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 63, 2, 237-264.
- Bénabou, R., (1996b), "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, 86, 3, 584-609.
- Bernard, A. and S. Durlauf, (1996), "Interpreting Tests of the Convergence Hypothesis," *Journal of Econometrics*, 71, 161-173.
- Billings, D., (1974), "Culture and Poverty in Appalachia: A Theoretical Discussion and Empirical Analysis," *Social Forces*, 53, 2, 315-323.

- Billings, D. and K. Blee, (2000), *The Road to Poverty*, New York: Cambridge University Press.
- Bollinger, C., J. Ziliak, and K. Trotske, (2009), "Down from the Mountain: Skill Upgrading and Wages in Appalachia," *IZA Discussion Paper* no. 2009.
- Brock, W. and S. Durlauf, (2001a), "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68, 2, 235-260.
- Brock, W. and S. Durlauf, (2001b), "Interactions-Based Models," in *Handbook of Econometrics*, Vol. 5, J. Heckman and E. Leamer, eds., Amsterdam: North Holland.
- Brock, W. and S. Durlauf, (2001c), "Growth Empirics and Reality," with W. Brock, *World Bank Economic Review*, 15, 229-272.
- Brock, W. and S. Durlauf, (2006), "Multinomial Choice with Social Interactions," with W. Brock, in *The Economy as an Evolving Complex System III*, L. Blume and S. Durlauf, eds., New York: Oxford University Press.
- Brock, W. and S. Durlauf, (2007), "Identification of Binary Choice Models with Social Interactions," with W. Brock, *Journal of Econometrics*, 140, 1, 52-75.
- Brock, W. and S. Durlauf, (2009), "Social Interactions and Adoption Curves," mimeo.
- Cochrane, J., (1988), "How Big Is the Random Walk in GNP?," *Journal of Political Economy*, 96, 5, 893-920.
- Cooper, S., D. Durlauf, and P. Johnson, (1994), "On the Evolution of Economic Status Across Generations," *American Statistical Association, Business and Economic Section, Papers and Proceedings*, 50-58, 1994.
- Cooley, J., (2008), "Desegregation and the Achievement Gap: Do Diverse Peers Help?" mimeo, University of Wisconsin at Madison.
- Crane, J., (1991), "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing," *American Journal of Sociology*, 96, 5, 1226-1259.
- Durlauf S., (1996a), "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1, 1, 75-93.
- Durlauf, S., (1996b), "Associational Redistribution: A Defense," *Politics and Society*, 24, 4, 391-410.
- Durlauf, S., (2004), "Neighborhood Effects," in *Handbook of Regional and Urban Economics*, vol. 4, J. V. Henderson and J.-F. Thisse, eds., Amsterdam: North Holland,
- Durlauf, S. and Y. Ioannides, (2009), "Social Interactions," mimeo.

Durlauf, S. and P. Johnson, (1995), "Multiple Regimes and Cross-Country Growth Behavior," with P. Johnson, *Journal of Applied Econometrics*, 10, 365-384.

Durlauf, S., P. Johnson and J. Temple (2005), "Growth Econometrics," in *Handbook of Economic Growth*, P. Aghion and S. Durlauf, eds., Amsterdam: North Holland.

Durlauf, S. and A. Seshadri, (2003), "Is Assortative Matching Efficient?," with A. Seshadri, *Economic Theory*, 21, 2-3, 475-493.

Eller, R., (2008), *Uneven Ground: Appalachia Since 1945*, Lexington: University of Kentucky Press.

Epple, D. and H. Sieg, (1999), "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, 107, 4, 645-681.

Evans, W., W. Oates, and R. Schwab, (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, 100, 5, 966-991.

Fryer, R. and P. Torelli, (2005), "An Empirical Analysis of Acting White," *NBER Working Paper no. 11334*.

Galor, O. and Zeira, J. (1993). "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 1, 35-53.

Glaeser, E., B. Sacerdote, and J. Scheinkman, (1996), "Crime and Social Interactions," *Quarterly Journal of Economics*, 111, 2, 507-548.

Graham. B., (2008), "Identifying Social Interactions through Conditional Variance Restrictions," *Econometrica*, 76, 3, 643-660.

Hanushek, E., J. Kain, M. Markman and S. Rivkin, (2003), "Does Peer Ability Affect Student Achievement?," *Journal of Applied Econometrics*, 18, 5, 527-544.

Heckman, J., (1979), Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153-61.

Heckman, J., (1998), "Detecting Discrimination," *Journal of Economic Perspectives*, 12, 2, 101-118.

Ioannides, Y. and J. Zabel, (2008), "Interactions, Neighborhood Selection and Housing Demand," *Journal of Urban Economics*, 63, 229-252.

Isserman, A., (1996), "Appalachia Then and Now: An Update of the "Realities of Deprivation" report to the President in 1964," mimeo, Appalachian Research Center.

Jencks, C. and S. Mayer, (1990), "The Social Consequences of Growing Up in a Poor Neighborhood," in *Inner-City Poverty in the United States*, L. Lynn and M. McGreary, eds., Washington D.C.: National Academy Press

Lichter, D. and L. Campbell, (2005), "Changing Patterns of Poverty and Spatial Inequality in Appalachia," mimeo, Appalachian Research Center.

Loury, G. (1977), "A Dynamic Theory of Racial Income Differences," in *Women, Minorities, and Employment Discrimination*, P. Wallace and A. Lamond, eds., Lexington: Lexington Books.

Loury, G., (1981), "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843-867.

Lucas, R. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22, 1, 3-42.

Lucas, R., (2009), "Ideas and Growth," *Economica*, 76, 301, 1-19.

Luhman, E., (1990), "Appalachian English Stereotypes: Language Attitudes in Kentucky," *Language in Society*, 19, 331-348.

Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60, 3, 531-542.

Manski, C., (2000), "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14, 3, 115-136.

Ogbu, J., (2003), *Black Students in an Affluent Suburb*, Mahwah NJ: Lawrence Erlbaum Associates.

Prat, A., (2002), "Should a Team be Homogeneous?," *European Economic Review*, 46, 1187-1207.

Romer, P. (1986), "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94, 5, 1002-1037.

Shaw, T., A. DeYoung, and E. Rademacher, (2004), "Educational Attainment in Appalachia: Growing with the Nation, But Challenges Remain," *Journal of Appalachian Studies*, 10, 3, 307-329.

Streufert, P., (2000), "The Effect of Underclass Isolation on School Choice," *Journal of Public Economic Theory*, 2, 4, 461-482.

Taylor, C., (1998), *Socrates*, New York: Oxford University Press.

Topa, G., (2001), "Social Interactions, Local Spillovers, and Unemployment," *Review of Economic Studies*, 68, 2, 261-295.

Wolfram, W. and D. Christian, (1976), *Appalachian Speech*, Arlington: Center for Applied Linguistics.

Wolfram, W. and N. Schilling-Estes, (2006), *American English*, New York: Basil Blackwell.

Yinger, J., (1995), *Closed Doors, Opportunities Lost*, New York: Russell Sage Foundation Press.