

ON THE EQUILIBRIUM DISTRIBUTION OF POPULATION
IN SPACE

MARTIN J. BECKMANN*

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY

Spatial equilibrium distributions of population are derived from the spatial distribution of net rates of reproduction, and from a relationship between migratory flow and gradients of population density and of locational "attractiveness." Conditions are discussed for which population approaches a uniform spatial density. Under certain conditions a particularly simple statement of the equilibrium conditions is possible in terms of the "potential of population," a concept introduced by demographers (J. Q. Stewart, *Geographical Review*, 37, 46-85, 1947) to measure the proximity of a point to people.

The problem of the distribution of population in space is of obvious interest and has attracted some attention in the literature. Empirical regularities concerning frequency distributions of city size originally discovered by G. Zipf (1941) have long posed a challenge to theoretical analysis. This has been met by N. Rashevsky (1947, pp. 93-107) and H. Simon (1955) among others. Another avenue of approach seemingly disconnected has been in terms of the potential of population, a concept modeled after Newtonian physics (Stewart, 1947). In this paper we propose to approach the equilibrium distribution of population in space in terms of a flow model of migration. With more specific assumptions this leads to a notion of the potential of population. No attempt will be made however to derive a frequency distribution of population density. It will appear that in our model such frequency distributions are ultimately due to the differential endowment of locations with economic opportunities, something we do not propose to explain here.

*This paper was first written at the Center for Advanced Study in the Behavioral Sciences.

Consider an economic region in a stationary state of equilibrium. At different locations, the densities of population and the per capita real incomes will, in general, be different. Partly as a result of this, partly for exogenous reasons, the reproductive rates of population will also differ within the region. In order to maintain the existing distribution of population and income, the surpluses and deficits in reproduction must be balanced by migration. Suppose that net immigration into the region as a whole is zero, so that we are concerned with internal migration only. To maintain a steady flow of migration, incentives must prevail in the form of interlocal differences in per capita real income or attractiveness of location. Thus equilibrium implies a complicated balance of incomes, population densities, reproduction rates, and migration flows. Under certain conditions the income level may tend to be uniform at all locations. But in general a more complicated state of equilibrium will result. The purpose of this paper is to develop the equilibrium conditions and solve them in the simpler cases.

1. *Preliminaries on Population Density and Income.* Both income and population density will appear as motivating forces in the generation of population deficits and surpluses and in the orientation of migration. But they are not independent variables. Given the economic opportunities at a location, the level of per capita income in real terms is a function of population density. To a smaller extent it also depends on the distribution of population densities outside the location considered. While we shall disregard the direct influence of these on the income level at a given location, it will appear that indirectly a balance exists between income level and the distribution of population around a given location.

The relationship between income level and population density at a given location may be assumed to follow the law of diminishing returns. That is, above a certain level per capita income is a decreasing function of population density. For simplicity, we shall sometimes use a linear approximation.

Let u, v be locational coordinates,
 $p(u, v)$ the population density,
 $y(u, v)$ the per capita (real) income level;

then the economic opportunities at a location are expressed by a function $f(p, u, v)$

$$y(u, v) = f[p(u, v), u, v], \quad (1.1)$$

which is decreasing with respect to p . Assuming f to be differentiable, we have

$$\frac{\partial f}{\partial p} < 0. \quad (1.2)$$

2. *Regeneration of Population.* What determines the surpluses and deficits of population *before* migration at different locations? We must distinguish for each location:

2.1 *The equilibrium rate c of reproduction.* This is the annual rate of reproduction required to sustain the given level of population in the reproductive age bracket. We shall disregard any disturbing effects of migration on the age composition of local populations. This is not a very unrealistic assumption as long as we restrict the notion of population to the reproductive ages. The equilibrium rate c depends only on mortality. Inasmuch as mortality differs among locations, c is a function of the locational coordinates

$$c = c(u, v).$$

To some extent mortality depends also on income levels and population densities. But we shall usually disregard this complication.

2.2 *The actual (net) rate of reproduction g .* This depends on various factors which we shall classify as:

population density	p
income level	y
location	u, v

The last is a catch-all for factors (such as attractiveness of the location) not absorbed in either population density (such as degree of urbanization) or income (such as economic opportunities).

Nothing can be said in general about the dependence of g on the locational coordinates. Of special interest is, of course, the case in which the rate of reproduction is uninfluenced by the location. Then the reproductive rate is the same function of population density and income at all locations.

$$g = g(p, y).$$

We consider next the relationship between rate of reproduction and population density. At very low densities of population any increase in density may be expected to raise the rate of reproduction

because of the increase in the number of contacts. From a certain level on, however, urbanization makes the rate of reproduction decrease with population density. As this effect grows stronger with rising population density, a point is finally reached from which on even the absolute level of reproduction (rate times density) decreases with any further increase in population density. We may think of this as a Malthusian equilibrium.

The effect of income on the reproduction rate is less transparent. A clear distinction must be drawn between the correlation of income and reproduction for a cross-section of a local population *at a given level* of average income, and the response of the aggregate rate of reproduction of this local population to a rise in average income. It is with the latter only that we are concerned here. There is evidence that "family size and income often tend to be positively associated within groups that are otherwise homogeneous" (Spengler, 1952, p. 101).

2.3 *Annual population surplus* e . This is the net of actual and equilibrium rate of reproduction multiplied by the population density. Substituting $y = f(p, u, v)$ from (1.1) we see that the rate of population surplus may be regarded as a function of population density and location alone:

$$e = e(p, u, v).$$

We shall be interested in Malthusian types of equilibria which are attained when population density pushes against the limits of resource availability. An increase of population density then decreases the annual population surplus either directly or indirectly through a fall in the income level.

3. *Forces Directing Migration*. We propose to consider migration as the net outcome of the random movements of many persons subject to an external field force. This force is assumed to be derived from a potential. Some alternative assumptions about its nature and composition will be examined.

In the simplest case, that is, in the absence of all external forces, a pure diffusion process* is present. Concentration itself

*An approach to human migration in terms of a diffusion process was made as early as 1921 by H. Hotelling (unpublished master's thesis, University of Washington, cited in Hotelling, 1949). The dynamics of the spatial expansion of biological populations has been analyzed in terms of random diffusion in an important paper by J. G. Skellam (1951).

assumes then the character of a repelling force. That is to say the vectors of net flow are everywhere a fixed proportion of the gradient of concentration. Population will thus diffuse from areas of high density into those of lower density. We denote by

$\varphi(u, v)$ the flow vector, whose direction is that of the flow and whose length equals the flow density

D the coefficient of diffusion
and obtain

$$\varphi(u, v) = -D \text{ grad } p(u, v). \tag{3.1}$$

In general φ is proportional to the vector sum of the diffusion force and the external force (Bjerknes, 1933, p. 117, equation 1). Let $a(u, v)$ be the potential of the external force. Then

$$\varphi(u, v) = \text{grad } a(u, v) - D \text{ grad } p(u, v). \tag{3.2}$$

The potential $a(u, v)$ itself may be interpreted as an index of the desirability of a location u, v . The assumption next in simplicity to $a(u, v) \equiv 0$ is that $a(u, v)$ be given a priori. It is then a function of location only. On the diffusion movements are now superimposed flows which are induced by the given differences in the attractiveness of the locations.

From an economic point of view it is more satisfactory to let the attractiveness of a location be influenced by its (real) income level, for prices and incomes are the forces by which demand is brought in line with supply. As applied to the present case this means that incomes play a compensatory part in attracting people to locations which are otherwise less desirable but in need of personnel. In general we shall assume therefore

$$a = a_0(u, v) + a_1[y(u, v)], \tag{3.3}$$

where a_1 is a given function, monotonically non-decreasing. In the simplest case the index of attractiveness may be so normalized that it is simply additive to income, resulting in a composite index of attractiveness

$$a = a_0(u, v) + y(u, v).$$

4. *The Equations of Equilibrium.* The only relationship needed to complete a system of equilibrium conditions is one equating the (annual) population net surplus with the (annual) rate of net emigra-

tion. This is supplied by the well-known source-sink equation or "equation of continuity" in hydrodynamics. (Kellog, 1929, p. 48)

$$\operatorname{div} \varphi(u, v) = q(u, v).$$

Here div denotes the divergence operator $\frac{\partial}{\partial u} + \frac{\partial}{\partial v}$, and $q(u, v)$ is the rate of net yield from sources or sinks located at (u, v) .

Taking the most general case we now have the following system of equations describing the equilibrium inside the region under consideration.

$$e(p, y, u, v) = \operatorname{div} \varphi(u, v), \quad (4.1)$$

$$y(u, v) = f(p, u, v), \quad (4.2)$$

$$\varphi(u, v) = -D \operatorname{grad} p(u, v) + \operatorname{grad} a(y, u, v). \quad (4.3)$$

To complete the picture, immigration into the region across the boundaries Γ must be specified. This means that the normal component φ_n of the flow vector must be given at all points of the boundary. For simplicity we shall assume that immigration is zero,

$$\varphi_n(u, v) = 0 \quad (u, v) \in \Gamma \quad (4.4)$$

The equation system may be reduced to a single equation for the interior of the region. First, y may be eliminated by substitution from (4.2). Next using the identity

$$\operatorname{div} \operatorname{grad} = \Delta = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}$$

we may combine (4.1) and (4.3):

$$-D \Delta p + \Delta a[f(p, u, v), u, v] = e[p, f(p, u, v), u, v]. \quad (4.5)$$

The boundary conditions also may be expressed in terms of p :

$$D(\operatorname{grad} p)_n = (\operatorname{grad} a[f(p, u, v), u, v])_n. \quad (4.6)$$

Solutions of this second-order partial differential equation of the boundary value type will now be discussed for alternative specifications of the data functions $a(y, u, v)$, $e(p, y, u, v)$, and $f(p, u, v)$.

5. *Homogeneous Region With No Migration at Boundaries.* Suppose that there are no inherent differences between locations of the region. That means that the locational coordinates as such do not

enter into the functions a, e, f . We will show that under mild assumptions the only equilibrium possible is one in which population is distributed at a uniform density and in which migration is absent.

The equations of equilibrium have the form

$$-D\Delta p + \Delta a[f(p)] = e[p, f(p)], \tag{5.1}$$

$$(-D \operatorname{grad} p + \operatorname{grad} a)_n = 0. \tag{5.2}$$

Write $-Dp + a = A$; let A_0 be an unspecified constant; and apply Green's First Identity (Kellog, *loc. cit.* p. 38) to

$$\iint (A - A_0) \Delta (A - A_0) \, dudv \quad \text{which yields} \tag{5.3}$$

$$= \oint (A - A_0) (\operatorname{grad} A)_n \, ds - \iint (\operatorname{grad} A)^2 \, dudv.$$

Because of the boundary condition, the first term on the right-hand side vanishes. Substituting $e[p, f(p)]$ for ΔA on the left-hand side we obtain

$$- \iint (\operatorname{grad} A)^2 \, dudv = \iint [A(p) - A_0] e[p, f(p)] \, dudv. \tag{5.4}$$

Now A is a decreasing function of p , since

$$\frac{dA}{dp} = -D + \frac{\partial a}{\partial y} \cdot \frac{df}{dp} \quad \text{and} \quad \frac{\partial a}{\partial y} > 0 \quad \text{by (3.3)}$$

$$\text{and} \quad \frac{df}{dp} < 0 \quad \text{by assumption (1.2).}$$

On the other hand $e[p, f(p)]$ is also decreasing, if we assume that either

$$\frac{\partial e}{\partial p} < 0 \quad \frac{\partial e}{\partial y} < 0 \quad \frac{df}{dp} < 0 \quad \text{or}$$

$$\frac{\partial e}{\partial p} > 0 \quad \frac{\partial e}{\partial y} > 0 \quad \frac{df}{dp} < 0$$

that is to say when we are at a point where the annual surplus of population decreases with population directly or at least indirectly because of the unfavorable effect on income of a population rise (cf. section 2.3). We now specify $A_0 = A(p_0)$ where $e[p_0, f(p_0)] = 0$. Then

$$[A(p) - A_0] e[p, f(p)] \geq 0, \tag{5.5}$$

for all p and “=” only if $p = p_0$. Since the right-hand side of (5.4) is therefore non-negative, and the left-hand side non-positive, we conclude that $p = p_0$.

We state as a conjecture here that the equilibrium is unstable under any alternative assumptions about e and f which would make $A(p)$ an increasing function. For this would permit an unchecked increase of population at any location.

6. *Linear Case.* Let the annual population surplus be linearly dependent on the population density at a location, with a uniform slope coefficient $-e_1$, e_1 being independent of position.

$$e[p, f(p), u, v] = e_0(u, v) - e_1 p. \quad (6.1)$$

This implies that the net rate of reproduction is a function of the form

$$g = \frac{g_0(u, v)}{p} - g_1,$$

where g_1 is a constant. For small densities of population, (in rural areas) the rate of reproduction is positive; for large densities it is negative. The equation of equilibrium has the form

$$-D\Delta p + \Delta a[f(p, u, v), u, v] = e_0(u, v) - e_1 p. \quad (6.2)$$

Consider now the special case of an unbounded region, whose population density vanishes at distances approaching infinity. For the following argument it is mathematically convenient to regard the distribution of population and the flow field as three dimensional, with coordinates u, v, w . Any limits one wishes to place on the extension into the third dimension are conveniently imposed through the densities a and e . The differential equation of population equilibrium remains unchanged if the operations div and grad are used in their three-dimensional meaning. The differential equation (6.2) may now be transformed like a Newtonian field equation (Kellog, *loc. cit.*, p. 156).

$$-Dp + a = \frac{1}{4\pi} \iiint \frac{e_0(U, V, W) - e_1 p(U, V, W)}{r(u, v, w, U, V, W)} dU dV dW, \quad (6.3)$$

where $r^2 = (u - U)^2 + (v - V)^2 + (w - W)^2$.

Write now

$$P(u, v, w) = \frac{1}{4\pi} \iiint \frac{p(U, V, W)}{r(u, v, w, U, V, W)} dU dV dW,$$

$$Q(u, v, w) = \frac{1}{4\pi} \iiint \frac{e_0(U, V, W)}{r(u, v, w, U, V, W)} dU dV dW.$$

The first expression is known in the literature as the potential of population. It is considered an inverse measure of the proximity of a point to people. A more ambitious interpretation regards it as the demographic counterpart of the Newtonian potential of mass attraction (Stewart, 1947). The second expression, which is not particularly noteworthy, is independent of the actual distribution of population and may therefore be considered a datum for each location.

The transformed equilibrium equation is thus

$$p(u, v, w) = \frac{1}{D} a(u, v, w) + \frac{e_1}{D} P(u, v, w) - \frac{1}{D} Q(u, v, w). \quad (6.4)$$

It states that under certain assumptions the equilibrium density of population is everywhere an increasing linear function of the opportunity density and the potential of population.

As Professor H. D. Landahl has pointed out to me, if e represents a "logistic" birth-death rate process so that $e = \alpha p - \beta p^2$, then we can write instead of (6.4)

$$p = \frac{a}{D} - \frac{\alpha}{D} P_1 + \frac{\beta}{D} P_2 - \frac{Q}{D},$$

where P_1 is a "first order" potential, P_2 is a "second order" or "second moment" potential such that if α and β depend on position, in general,

$$P_1 \bar{\alpha} = \frac{1}{4\pi} \iiint \frac{\alpha(U, V, W) p dU dV dW}{r},$$

$$P_2 \bar{\beta} = \frac{1}{4\pi} \iiint \frac{\beta(U, V, W) p^2 dU dV dW}{r}.$$

The empirical observation that in *rural areas* population density appears to be proportional to the square of the potential of population (Stewart, *loc. cit.*) rather than being linearly related to it suggests that the mobility of rural population is different from that suggested by a pure diffusion model. It is tempting to introduce differential degrees of mobility into the model by letting flow be proportional to the gradient of some power, with exponent less than unity, of population density. An exponent of $\frac{1}{2}$ would be compatible with the observed rural density distribution, provided that a and Q would roughly cancel out. However, such special hypotheses shall not be further pursued here.

In revising this paper, I have benefited from penetrating comments by Professors N. Rashevsky and H. D. Landahl, for which I express my gratitude.

LITERATURE

- Bjerknes, V., J. Bjerknes, H. Solberg, and T. Bergenson. 1933. *Physikalische Hydrodynamik*. Berlin: J. Springer.
- Hotelling, H. 1949. "Stochastic Processes, Historical Summary of the Problem." (abstract) *Econometrica*, 17, 66-8.
- Kellog, O. D. 1929. *Foundations of Potential Theory*. F. Ungar: New York.
- Rashevsky, N. 1947. *Mathematical Theory of Human Relations*, Bloomington, Indiana: Principia Press.
- Simon, H. 1955. "On a Class of Skew Distribution Functions." *Biometrika*, 42, Parts 3 and 4, 425-40.
- Skellam, J. G. 1951. "Random Dispersal in Theoretical Populations." *Biometrika*, 38, 196-217.
- Spengler, J. J. 1952. "Population Theory." *A Survey of Contemporary Economics*, Vol. 2, pp. 83-131. (B. F. Haley, Ed.) Homewood, Ill.: R. D. Irwin.
- Stewart, J. Q. 1947. "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population." *Geographical Review*, 37, 46-85.
- Zipf, G. K. 1941. *National Unity and Disunity*. Bloomington, Indiana: Principia Press.

RECEIVED 12-19-56