POPULATION MOMENTUM FOR GRADUAL DEMOGRAPHIC TRANSITIONS: AN ALTERNATIVE APPROACH*

JOSHUA R. GOLDSTEIN

In this article, I derive a simple formula for approximating the ultimate size of a population that undergoes a gradual transition to replacement fertility. I model the fertility transition by specifying a linear frontier on the Lexis surface across which a change in fertility is instantaneous. Gradual transitions result from variations in the slope of this frontier. This framework can be used to reproduce and understand previous studies of population momentum and gradual transitions.

In an influential article, Keyfitz (1971) presented a simple, yet powerful scenario for illustrating the inevitability of population growth even after the fertility rate falls to its replacement level. Even if the decline in fertility is immediate, the population continues to grow because of its relatively young age structure. Keyfitz presented a simple formula for estimating the population momentum implied by the age structure of growing populations. Today his scenario and the results of his formula are included in the most influential world population forecasts as a benchmark against which to compare the projected future evolution of population sizes (Bos et al. 1992; Bos et al. 1994; United Nations 1998).

In practice, as Keyfitz observed, the inevitability of continued population growth comes not only from the current age structure but also from the time it takes for fertility rates to fall to replacement level. Replacement fertility is typically reached only after decades of decline, not overnight. Thus, population researchers have long sought a formula for calculating the effect of more realistic gradual demographic transitions. As recently as 1998, Schoen and Kim wrote, "There is no analytical way to determine what additional growth will result from a gradual transition to replacement level rates" (p. 295). They instead showed that such a transition in rates could be modeled by assuming a gradual leveling of the stream of births. Subsequently, however, Li and Tuljapurkar (1999, 2000) presented general results for changing demographic rates. Using mathematical results from renewal theory, they derived the momentum implied by a range of speeds and forms of fertility decline. In particular, they found a simple formula for determining the effects of a gradual linear decline in net fertility.

Here, I present an alternative derivation of some of Li and Tuljapurkar's important results in a less general but simpler form. Instead of their period approach to fertility decline, I use a framework closely related to the cohort-based scenario introduced by Frauenthal (1975). Whereas the period approach produces a simple, exact formula only for rapid transitions (shorter than about 15 years), the cohort approach used here is exact for transitions that last about twice as long (some 30 to 35 years). The approach given here has the additional advantage of including not just Keyfitz's (1971) formula as a special case but also that of Frauenthal. It helps provide intuition for Li and

*Joshua R. Goldstein, Office of Population Research, Princeton University, Wallace Hall, Princeton, NJ 08544; E-mail: josh@princeton.edu. I thank Henry Goldstein, Nan Li, Michael McKenna, Arodys Robles, Norm Ryder, Wilhelm Schlag, John Stalker, Guy Stecklov, James Trussell, Shripad Tuljapurkar, Kenneth Wachter, and Charles Westoff for their help and comments. I am solely responsible for any errors.

Tuljapurkar's results. Finally, it also is consistent with Schoen and Kim's (1998) estimates for population momentum resulting from a gradual leveling of the stream of births.

The consistency of the various alternative formulations points to the means by which a simple lesson can be learned from the complicated dynamics of population momentum for gradual transitions. What is important for the eventual population size is how far fertility rates fall and how long the transition lasts. If forecasters can get these two factors correct, then the details of the time and age patterns of the fertility transition will make relatively little difference.

A LEXIS FRONTIER APPROACH TO MODELING DEMOGRAPHIC TRANSITIONS

I take as my starting point a population that has a long history of constant age-specific fertility and mortality rates. I then specify a frontier on the age-time plane of the Lexis diagram. As a cohort crosses this frontier, it undergoes an immediate transition to replacement-level age-specific fertility rates. Tilting the frontier translates the instant transition into a gradual transition with a range of speeds.

In theory, this frontier could take a wide variety of forms, but here I consider only frontiers that are straight lines. This simple approach is flexible enough to duplicate or approximate a wide range of scenarios in the literature. Examples of linear transition frontiers are illustrated in panels a, b, and c of Figure 1. To the left of the frontiers the populations are subject to the net fertility rates ϕ_1 of a stable population. To the right, they are subject to the rates ϕ_2 of a stationary population.

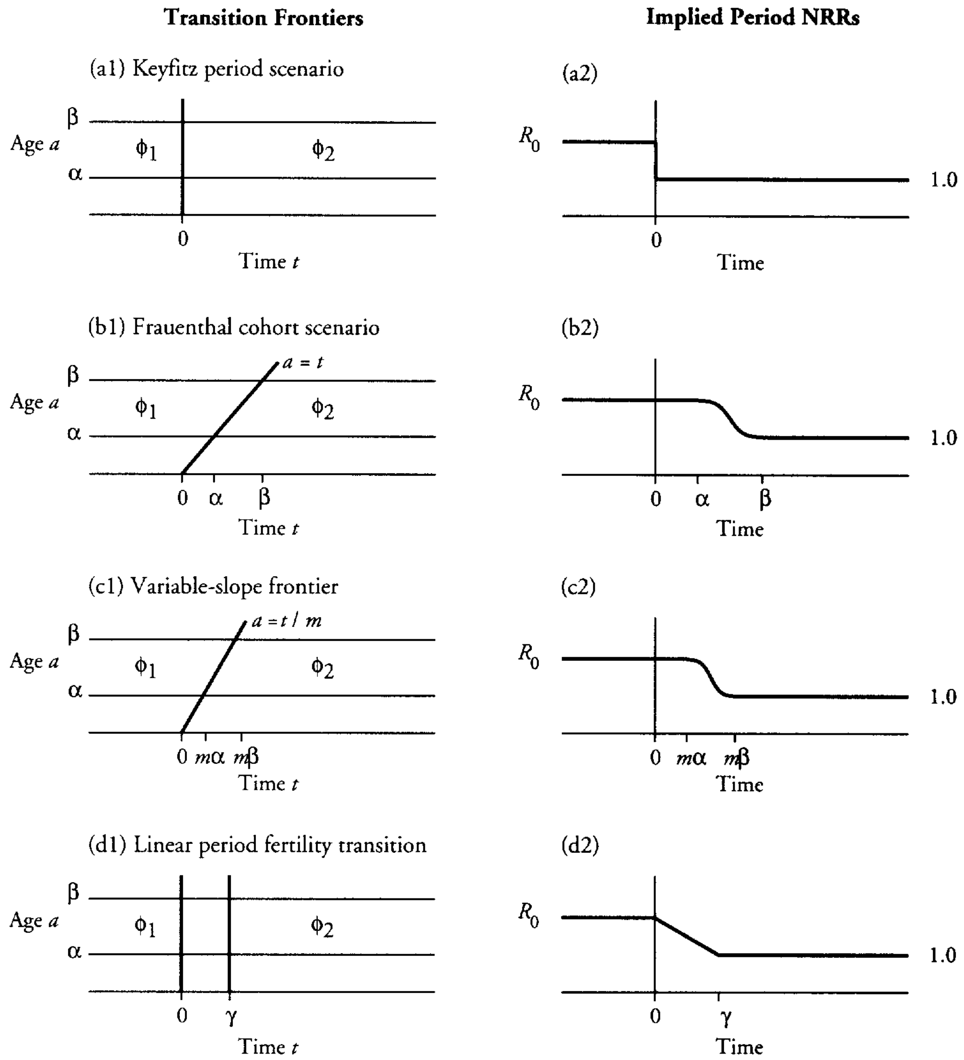
Keyfitz's classic scenario is shown in panel a1. In panel a2, a vertical frontier describes an instantaneous drop in fertility from one period to the next. Frauenthal (1975) proposed that fertility might drop to replacement from one cohort to the next (see panel b1): all individuals who were born before time zero would be subject to the old net maternity function for their entire lives, and all individuals who were born after time zero would be subject to the new net maternity function for their entire lives. This scenario corresponds to a frontier angled at 45 degrees, running along the line $a = t$. Frauenthal's scenario, even though it involves a transition that takes place instantly from one cohort to the next, corresponds to a *gradual* S-shaped change in net period fertility (see panel b2). Both Keyfitz's and Frauenthal's scenarios are specific cases of the general variable-slope fertility transition frontier shown in panel c1.

The variable frontier approach has both a literal and a more useful, but indirect, interpretation. If interpreted strictly, the frontier marks the point in time when the fertility of a given age group suddenly falls. A line angled between 45 and 90 degrees would correspond to a fertility decline that begins at the youngest age groups and smoothly spreads to adjacent ages as the population moves through time. A line angled at more than 90 degrees (not shown) would describe a population in which replacement fertility begins at the oldest ages and spreads to the youngest. The variable-slope frontier thus provides a model for the age-based diffusion of fertility control.

An alternative way to think of the Lexis frontier model is as a simple way to model changes in period fertility. The length of the period fertility transition can range from 0 years (when the slope is infinite) to the span of a cohort's reproductive interval, some 30 to 35 years. This formulation allowed me to model long-term population size using a single parameter, the slope of the frontier.

The linear variable-slope frontier on the Lexis diagram produces a logistic-like decline in the period net reproduction rate (NRR; see panel c2) because human net maternity functions tend to be bell shaped. The smooth decline in period fertility resulting from the variable-slope frontier can be compared with the sharp corners of the linear period fertility transition modeled by Li and Tuljapurkar (1999). The Lexis diagram of the linear

Figure 1. Lexis Diagrams of Transition Frontiers of Instantaneous and Gradual Fertility Transitions and Typical Implied Trajectories of the Net Reproduction Rate (NRR)



Notes: Before crossing the linear frontiers on the Lexis diagrams, populations are stable, growing at the intrinsic rate determined by the net maternity function ϕ_1. After the transition, fertility is governed by the replacement level net maternity function ϕ_2. The Keyfitz (1971) scenario is depicted in panels a1 and a2. Frauenthal's (1975) cohort scenario is depicted in panels b1 and b2. The variable-slope frontier scenario used in this article is shown in panels c1 and c2. Li and Tuljapurkar's (1999, 2000) linear period decline is shown in panels d1 and d2.

period fertility transition is in panel d1, where the two vertical lines at time 0 and time γ mark the interval over which the period net maternity function changes linearly from the old to the new regime.

FORMULAS FOR FERTILITY TRANSITIONS

With this background, we now turn to the mathematical results describing the effect of the fertility transition frontier on the ultimate size of the stationary birth stream and population size.

I begin with the vertical case (panel a1), where the equation for the vertical line is simply $t = 0$. Using different techniques from those that are used here, Keyfitz (1971) found the birth stream $B(t)$ converges to a constant $B(\infty)$ such that

$$\frac{B(\infty)}{B(0)} \approx \frac{1}{r\mu}\left(\frac{R_0 - 1}{R_0}\right), \tag{1}$$

where r is the growth rate of the stable population prior to the transition, μ is the mean age of childbearing (the same in both the initial and ultimate populations), and R_0 is the NRR in the initial population. The amount of population momentum, that is, the ratio of the size of the ultimate stationary population to that of the initial stable population, can be found by multiplying either side by the product of the expectation of life at birth e_0 and the birth rate b_0 of the initial population.

Frauenthal (1975) extended Keyfitz's result to cover a particular kind of gradual demographic transition. Letting cohort-fertility drop to replacement instantaneously from one cohort to another produces a gradual decline over a period equal to the span of the reproductive years $\beta - \alpha$ (see panels b1 and b2 of Figure 1). Frauenthal's scenario corresponds to a transition path along the line $a = t$. Frauenthal found that in this case

$$\frac{B(\infty)}{B(0)} \approx \frac{1}{r\mu}\left(R_0 - 1\right). \tag{2}$$

Frauenthal's result is exactly a factor of R_0 greater than Keyfitz's. Shifting the frontier to the cohort line delays the transition by a different amount for each age group. On average across all age groups, however, the high-fertility regime persists about the mean age of childbearing μ years longer. The Frauenthal scenario produces birth cohorts that, upon transition, are about $e^{r\mu} \approx R_0$ times larger than they would have been.

I now generalize to a frontier that can take a range of slopes across the Lexis surface. This case is illustrated in panels c1 and c2 of Figure 1. The linear frontier produces a gradual, logistic-shaped period transition. Let $1 / m$ be the slope of the linear frontier $a = t/m$; $m = 0$ corresponds to the vertical frontier, and $m = 1$ corresponds to the cohort frontier. Then, for $m \geq 0$, I find

$$\frac{B(\infty)}{B(0)} \approx \frac{1}{r\mu}\left(\frac{R_0 - 1}{R_0}\right)\left(R_0\right)^m. \tag{3}$$

This equation is exact when $m = 0$ and $m = 1$.

This result can be understood intuitively by noting that as in the Frauenthal case, the variable-slope linear frontier involves delaying the transition by different amounts of time for different ages. Averaging this delay over all ages is like having an instantaneous transition for all ages some $m\mu$ years after time 0. By this time, the number of women giving birth is a factor $e^{rm\mu} \approx (R_0)^m$ times larger than it would have been in the Keyfitz scenario.

Both Keyfitz's and Frauenthal's scenarios emerge as special cases of this formula (by letting m equal 0 and 1, respectively). In addition, this formula is approximately the same as Li and Tuljapurkar's result for gradual demographic transitions that involve a linear pattern of period fertility decline (see panels d1 and d2 of Figure 1).

Although the following derivation is given only for forward-sloping frontiers, it also applies to backward-sloping frontiers that involve rapid transitions. The transition must be short enough so that the children who are born after the transition do not themselves cross the frontier a second time. For values of $m > 1$, a similar issue arises. The daughter or, for a very large m, even the granddaughter generations can end up recrossing the frontier. Li and Tuljapurkar (1999) incorporated multiple generations into the more complex versions of their model. They found, however, that the single-generation formula is generally a reliable approximation for longer transitions.

DERIVATION OF MOMENTUM FORMULA

I derive Eq. (3) by taking advantage of a basic result in renewal theory. For $0 \le m \le 1$, all cohorts born after time $t = 0$ are subject to the stationary net maternity function $\phi_2(a)$. Accordingly, the renewal equation following the transition will take the form

$$B(t) = \int_0^t \phi_2(a)B(t-a)da + g_0(t), \quad t > 0 \tag{4}$$

where $B(t)$ is the stream of births at time t, $\phi_2(a)$ is the net maternity function at age a, and $g_0(t)$ is the number of births produced by survivors of the initial population present at time $t = 0$.

The basic renewal theorem indicates that in the long run, the stream of births subject to a constant net maternity function will converge to a limit (e.g., Feller 1971). Specifically, if $\int_0^\infty \phi_2(a)da = 1$, which it does by definition because eventual fertility after the transition is at replacement levels,

$$\lim_{t\to\infty} B(t) = \frac{1}{\mu}\int_0^\infty g_0(t)dt, \tag{5}$$

where μ is the mean age at birth in the stationary population.

Determining the size of the eventual birth stream thus involves calculating the sum of $g_0(t)$, the births produced by survivors of the initial population. Births occurring after time 0 to women born before time 0 can consist of children born under the old regime of fertility ϕ_1 and of children born under the new regime of fertility ϕ_2. Specifically,

$$g_0(t) = \int_t^{t/m} B(t-a)\phi_2(a)da + \int_{t/m}^\infty B(t-a)\phi_1(a)da.$$

When $m = 1$, no such children are born under the new regime; when $m = 0$, no such children are born under the old regime.

Integrating $g_0(t)$ involves reversing the order of integration and then substituting the stable historical birth stream $B(0)e^{-r(t-a)}$ for $B(t-a)$. For simplicity, let $B(0) = 1$ to obtain the general expression

$$\int_0^\infty g_0(t)dt = \frac{1}{r}\left[\int_0^\infty e^{r(m-1)a}\phi_1(a)da - \int_0^\infty e^{r(m-1)a}\phi_2(a)da\right]. \tag{6}$$

This expression is exact and applies to any replacement level net-maternity function ϕ_2. One can obtain a simpler expression by assuming that fertility declines by the same proportion at all ages (i.e., $\phi_2(a) = \phi_1(a) / \int \phi_1(a)da$). This is the assumption used by Keyfitz (1971) and Frauenthal (1975).[1] Under proportional fertility declines at all ages, Eq. (6) simplifies to

$$\int_0^\infty g_0(t)dt = \frac{1}{r}\left(1 - \frac{1}{R_0}\right)\int_0^\infty e^{r(m-1)a}\phi_1(a)da. \tag{7}$$

1. See Mitra (1976) for an exploration of what happens when this assumption is violated.

The right-hand integral corresponds to the integral of Lotka's equation, except that the net maternity function is discounted by $e^{-ra} \times e^{rma}$ instead of simply by simply by e^{-ra}.

One can approximate the integral by its cumulant expansion (Kendall 1952: chap. 5) about $r = 0$. Let

$$Y(r) = \log \int_0^\infty e^{r(m-1)a} \phi_1(a) da. \tag{8}$$

The first three terms are

$$Y(r) \approx \log(R_0) + (m-1)\mu r + \frac{1}{2}(m-1)^2 \sigma^2 r^2, \tag{9}$$

where μ is the mean age of childbearing in the stable population with growth rate r; σ^2 is the variance of the net-maternity function in the stable population. In human populations, $\sigma^2 r^2$ is small compared with μr, making the contribution of the third term negligible. When the first two terms are employed and the standard approximation, $R_0 \approx e^{\mu r}$, is used, $e^{Y(r)} \approx R_0 \times (R_0)^{m-1} = (R_0)^m$. Incorporating this approximation into Eq. (7) produces the general result given in Eq. (3).

Frauenthal (1975) noted that the factor $(R_0 - 1) / (r\mu R_0)$ is well approximated by $(R_0)^{-1/2}$. This factor appears in Keyfitz's (1971) scenario and in my Eq. (3). Using this additional approximation produces an even simpler expression for the eventual birth stream under gradual transitions:

$$\frac{B(\infty)}{B(0)} \approx (R_0)^{m-(1/2)}. \tag{10}$$

Let b be the crude birth rate at the time the transition begins and e_0 be life expectancy. The factor by which total population size will eventually increase as a result of the transition then will be

$$M \approx b e_0 (R_0)^{m-(1/2)}. \tag{11}$$

CORRESPONDENCE BETWEEN THE FORMULAS

As I noted earlier, the Keyfitz (1971) and Frauenthal (1975) scenarios appear as immediate cases of my general formula. I now show that my approach is consistent with earlier results given by Li and Tuljapurkar (1999) and Schoen and Kim (1998).

I found that a variable-slope frontier will increase the momentum of fertility transitions by a factor of $(R_0)^m$ when compared with Keyfitz's instant transition. Li and Tuljapurkar showed the extra momentum of a transition in which fertility at all ages declines linearly over a period of γ years is equal to $(e^{r\gamma} - 1) / (r\gamma)$. Because $(e^x - 1) / x \approx e^{x/2}$, Li and Tuljapurkar's result can be rewritten as approximately equal to $\exp(r\gamma / 2)$. This last approximation is the same as the result for extra momentum that Schoen and Kim (1998) found in their study of the special case in which the NRR changes in such a manner as to cause the growth rate of the birth stream itself to decline linearly.

This approximation corresponds to my result in the case where $m = 1$. The Frauenthal scenario of instant cohort fertility transition generates a gradual change in the period NRR (e.g., panels b1 and b2 in Figure 1). The original NRR level continues for duration α, then declines gradually to replacement level over duration $(\beta - \alpha)$, where α and β are the beginning and end, respectively, of childbearing ages. The transition ends after β years, when the last members of the original cohort finish childbearing. The extra population growth during this transition is composed of two parts: $e^{(r\alpha)}$ during the first a years, and—following either Li and Tuljapurkar or Schoen and Kim—$e^{(r(\beta - \alpha)/2)}$ during the subsequent

($\beta - \alpha$) years, where r is the growth rates of the initial stable population. If the human fertility pattern is such that $\alpha \approx (\beta - \alpha) / 2 \approx \mu/2$, where μ is the mean age of childbearing, the two growth factors become $e^{(r\alpha)} \approx e^{(r\mu/2)} \approx \sqrt{R_0}$, and $e^{(r(\beta - \alpha)/2)} \approx e^{(r\mu/2)} \approx \sqrt{R_0}$, respectively. Their product is then R_0, the NRR of the initial population.

For $m \neq 1$, the above correspondence extends naturally (see panels c1 and c2 of Figure 1); that is, the growth during the first $m\alpha$ years is $e^{(rm\alpha)}$, and growth during the subsequent ($m\beta - m\alpha$) years is $e^{(rm(\beta - \alpha)/2)}$. Again, if the human fertility pattern is such that $\alpha \approx (\beta - \alpha)/2 \approx \mu/2$, then the product of these two factors yields $e^{(rm\mu)} \approx (R_0)^m$. The equivalence between the various formulations is not exact. It also is subject to the condition that mean age of childbearing is centered within the reproductive span. However, because typical values in human populations are $\alpha \approx 15$, $\beta \approx 45$, and $\mu \approx 30$, the equivalence of the two formulations is fairly general.

As a final check on the accuracy of the approximations used both to derive the simple gradual momentum formula for linear transition frontiers and to show the correspondence between them, estimates of momentum were made numerically according to the various approaches, for $0 \leq m \leq 1$. Differences should be the highest for fast-growing populations, so I used 1983 period rates for Mexico: NRR = 2.141, $r = 0.027$ (Keyfitz and Flieger 1990). Using Eq. (3), I found that population momentum would increase population size by factors ranging from 1.70 when $m = 0$ to 3.64 when $m = 1$. The exact numerical integration of Eq. (7) produced estimates of population momentum that differ from the approximate result in Eq. (3) by a maximum of about half a percentage point over the interval $0 < m < 1$ and were exact for $m = 0$ and $m = 1$. Comparing Eq. (3) with Li and Tuljapurkar's results using the Mexican data produced a maximum discrepancy of less than 5%. Compared with the approximation of Li and Tuljapurkar's results (or, equivalently, Schoen and Kim's), Eq. (3) estimates differ by, at most, 2%.

One would not expect perfect equality among the various formulations because the linear frontier transition does not, in general, produce a linear period transition.[2] In human populations, the net maternity function is roughly bell shaped, so a sudden transition across a linear frontier produces a backward S-shaped, approximately logistic, pattern of decline. Li and Tuljapurkar found in their simulations that the logistic and linear declines produce essentially the same long-run consequences for population size.

The correspondence between all three approaches shown here is additional evidence that what really matters in determining population momentum in gradual demographic transitions is simply the magnitude and speed of the transitions. Other details, like the age pattern of fertility and the precise trajectory of vital rates—dimensions across which the various approaches do, indeed, differ—do not seem to matter much.

DISCUSSION

This article has presented an alternative approach for estimating the long-run size of populations that undergo transitions to replacement fertility. I introduced a method for varying the speed of fertility transitions by changing the slope of a linear frontier on the Lexis surface. On the basis of this variable frontier model, I derived a formula for population momentum that agrees with the result found by Li and Tuljapurkar (1999, 2000) for linear period fertility transitions. The agreement between the two approaches reinforces the finding that the overall level of fertility and the speed of its decline play leading roles in determining the long-run size of populations approaching stationarity. Changes in the shape of the fertility age schedule play a minor role, as long as the mean age of childbearing remains constant.[3]

2. A linear period transition would occur only if the net maternity function were rectangular.

3. Changes in the mean age of childbearing, however, can have large effects on population size. See Mitra (1976) and Goldstein and Schlag (1999).

Table 1. Ultimate Population Sizes for Pakistan, by Speed of Fertility Decline and Life Expectancy

Speed of Fertility Decline	m	Year of Completed Transition	e_0	Ultimate Population Size (in Millions)	Growth Factor
Instant	0	1990	56	170	1.52
Fast	1/2	2015	56	251	2.24
Slow	1	2040	56	370	3.30
Instant	0	1990	85	259	2.31
Fast	1/2	2015	85	381	3.40
Slow	1	2040	85	561	5.00

The speed at which fertility declines can make an enormous difference in eventual population size. Take, for example, Pakistan, which in 1990 had a population of 112 million people, an NRR of 2.17, a crude birth rate of about 40 per thousand, and a life expectancy at birth of about 56 years. Table 1 shows the results obtained by applying Eq. (11) to the case of Pakistan. As can be seen, a gradual transition lasting about 50 years produces an ultimate population that is roughly twice as large as that implied by an instant transition (a population size of 370 million instead of 170 million).

It is possible to extend the applicability of results for gradual momentum by incorporating increases in longevity. Mortality declines at ages beyond the maximum age of childbearing will enter into ultimate population size via the expectation-of-life term (e_0) used in the momentum formula (Goldstein and Stecklov 2001; Ryder 1975). Accounting for an increase in longevity forecast by the World Bank (e_0 = 85 years by 2150), the population size implied by a "fast" transition rises to 381 million and that implied by a "slow" transition, to 561 million.

These figures are in line with the World Bank's actual long-term forecast of 397 million for Pakistan in 2150. The World Bank used a full cohort-component computer projection to obtain its forecasts. Agencies may want to consider using analytic formulas to replace or supplement computer projections. The formulas are quicker and easier to use than the computer projection, and have the added advantage that the forecaster can see immediately the impact of various assumptions about the speed of demographic transitions. The gradual momentum results can be extended to transitions to nonstationary stable states and can produce results that are close to the high, low, and medium long-term forecasts made by the United Nations (Goldstein and Stecklov 2001).

The examples and discussion in this article, as well as in the articles by Keyfitz (1971), Frauenthal (1975), Mitra (1976), Schoen and Kim (1998), and Li and Tuljapurkar (1999, 2000) have focused on populations with histories of positive growth that undergo fertility declines. The mathematics, however, applies to all growth rates. In particular, it applies to shrinking populations, for which population momentum implies further declines in population size, even after fertility rises to replacement.

REFERENCES

Bos, E., M.T. Vu, A. Levin, and R.A. Bulatao. 1992. *World Population Projections, 1992–1993.* Baltimore: John Hopkins University Press.

Bos, E., M.T. Vu, E. Massiah, and R.A. Bulatao. 1994. *World Population Projections 1994–95: Estimates and Projections With Related Demographic Statistics.* Baltimore: Johns Hopkins University Press.

Feller, W. 1971. *An Introduction to Probability Theory and Its Applications,* Vol. 2. 2nd ed. New York: John Wiley and Sons.

Frauenthal, J.C. 1975. "Birth Trajectory Under Changing Fertility Conditions." *Demography* 12:447–54.

Goldstein, J.R. and W. Schlag. 1999. "Longer Life and Population Growth." *Population and Development Review* 25:741–47.

Goldstein, J.R. and G. Stecklov. 2001. "The Virtues of Simplicity: Long Run Population Forecasting Without Computers." Unpublished manuscript. Office of Population Research, Princeton University.

Kendall, M.G. 1952. *The Advanced Theory of Statistics.* 5th ed. New York: Hafner Publishing.

Keyfitz, N. 1971. "On the Momentum of Population Growth." *Demography* 8:71–80.

Keyfitz, N. and W. Flieger. 1990. *World Population Growth and Aging.* Chicago: University of Chicago Press.

Li, N. and S. Tuljapurkar. 1999. "Population Momentum for Gradual Demographic Transitions." *Population Studies* 53:255–62.

———. 2000. "The Solution of Time-Dependent Population Models." *Mathematical Population Studies* 74:311–29.

Mitra, S. 1976. "Influence of Instantaneous Fertility Decline to Replacement Level on Population Growth: An Alternative Model." *Demography* 13:513–19.

Ryder, N. 1975. "Notes on Stationary Populations." *Population Index* 41(1):3–28.

Schoen, R. and Y.J. Kim. 1998. "Momentum Under a Gradual Approach to Zero Growth." *Population Studies* 52:295–99.

United Nations, Department of International and Economic and Social Affairs. 1992. *Long-Range World Population Projections: Two Centuries of Population Growth 1950–2150.* New York: United Nations.