International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Predictability, Complexity, and Catastrophe in a Collapsible Model of Population, Development, and Environmental Interactions

Sanderson, W.C.

IIASA Working Paper

WP-94-075

August 1994

Working Paper

Predictability, Complexity, and Catastrophe in a Collapsible Model of Population, Development, and Environmental Interactions

Warren C. Sanderson

WP-94-75
August 1994

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg Austria

Telephone: +43 2236 715210 □ Telex: 079137 iiasa a □ Telefax: +43 2236 71313

Predictability, Complexity, and Catastrophe in a Collapsible Model of Population, Development, and Environmental Interactions

Warren C. Sanderson

WP-94-75
August 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg Austria
Telephone: +43 2236 715210 □ Telex: 079137 iiasa a □ Telefax: +43 2236 71313

ABOUT THE AUTHOR

Professor Warren C. Sanderson is from the Department of Economics, State University of New York at Stony Brook, USA. During June-August 1994, he visited IIASA as a Research Scholar.

ABSTRACT

More and more population forecasts are being produced with associated 95 percent confidence intervals. How confident are we of those confidence intervals? In this paper, we produce a simulated dataset in which we know both past and future population sizes, and the true 95 percent confidence intervals at various future dates. We use the past data to produce population forecasts and estimated 95 percent confidence intervals using various methodologies. We, then, compare the true 95 percent confidence intervals with the estimated ones. This comparison shows that we are not at all confident in the estimated 95 percent confidence intervals. Indeed, even calling them confidence intervals is quite misleading.

TABLE OF CONTENTS

PREDICTABILITY, COMPLEXITY, AND CATASTROPHE IN A COLLAPSIBLE MODEL OF POPULATION, DEVELOPMENT, AND ENVIRONMENTAL INTERACTIONS

Warren C. Sanderson

"Demographers can no more be held responsible for inaccuracy in forecasting 20 years ahead than geologists, meteorologists, or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be." Keyfitz (1981, p. 579)

1. Introduction

More and more population forecasts are being produced with associated 95 percent confidence intervals.[1] How confident are we of those confidence intervals? In this paper, we produce a simulated dataset in which we know both past and future population sizes, and the true 95 percent confidence intervals at various future dates. We use the past data to produce population forecasts and estimated 95 percent confidence intervals using various methodologies. We, then, compare the true 95 percent confidence intervals with the estimated ones. This comparison shows that we are not at all confident in the estimated 95 percent confidence intervals. Indeed, even calling them confidence intervals is quite misleading.

Section 2 contains a discussion of the nature of the confidence intervals that we study here. The model that produces past and future populations and associated true confidence intervals is discussed in Section 3. Our model is known as Wonderland,[2] and it is a simple place where we can see the interactions of population growth, economic growth, and environmental change. Section 4 presents the "true" 95 percent confidence intervals for Wonderland's future population.

In Section 5, we use a variety of approaches, based on Wonderland data from 1930 to 1995, to produce expected populations and confidence intervals for the years 2025, 2050, 2075, and 2100. It is in this section that we see that the computed confidence intervals can be quite different from the true ones. In Section 6, we use data from 1855 through 1995 and focus on the expected populations and confidence intervals for 2005, 2015, and 2025. We provide an example of circumstances under which the true and computed confidence intervals are disjoint, even in the short-run. Section 7 concludes with some suggestions to forecasters, based on what we have learned.

---

[1]A highly incomplete list of demographic papers that use 95 percent confidence intervals can be found in the References.

[2]An earlier non-stochastic version of the Wonderland model appears in Sanderson (1994).

2. What Are Future Confidence Intervals?

We distinguish two views of the future here, one that the future is deterministic and the other that the future is stochastic. If the future were deterministic, the best population forecast would have zero variance. It would simply be the true future population size. In that case, uncertainty about the future is a measure of the ignorance of the researcher and should be minimized. If the future were stochastic, then forecasters need to produce estimates of the distribution of future population sizes. Uncertainty, in that case, would be part of nature. A forecast with the correct expected population size and a zero variance could be a very inaccurate one. We take the position here that nature is stochastic and that the 95 percent confidence intervals that forecasters are trying to measure are the intervals within which future population sizes will fall 95 percent of the time.

3. Wonderland

Wonderland is a simple nine equation stochastic model of economic, demographic, and environmental interactions. Everything that is needed to generate historical and future data for Wonderland, the equations, parameters, and initial conditions, are given in Appendix A. In Wonderland, we have at our disposal the time paths of population, the crude birth rate, the crude death rate, real per capita GDP, the annual flow of pollution, pollution control expenditures and the stock of natural capital. The stock of natural capital may be thought of as the set of things provided to us by the environment, like air and water, which allow us to live healthy and productive lives.

Reality in Wonderland is uncertain in two different respects. Given the parameters, the evolution of Wonderland depends on six random numbers in each year. The distributions of the random variables that generate those numbers are assumed to be known and the same in the future as in the past. There is also uncertainty in Wonderland because an important parameter is unknown, the one that influences how much pollution will be produced per unit of output in the future. For simplicity, we assume that it can take on only one of two possible values. On one side of the looking-glass, we assume a value that generates a relatively high amount of pollution. We call the situation on this side of the looking-glass the Environmentalists' Nightmare Scenario (ENS) for reasons that will become apparent in just a moment. On the other side of the looking-glass, everything is identical except that we assume a value of the parameter that produces less pollution. We call this situation the Economists' Dream Scenario, for reasons that will also become apparent in a moment.[3]

At first glance, Wonderland might look like an unusual place in which to study population projections because it has no age structure and no migration. Yet, because of its simplicity, lessons will be easier to draw there and they will apply to populations with age structures and migration flows as well.

Figures 1 through 5 show one particular stochastic history and future of Wonderland, based on the information in Appendix A. The time series of Wonderland's population appears in Figure 1. Even though the history and future are stochastic, the time series of population in

[3]To be precise, we assume that the value of the χ parameter in equation (A9) is -0.01 in the environmentalists' nightmare scenario, and -0.03 in the economists' dream scenario. Otherwise, everything is identical across the two scenarios. The implications of this parameter choice for pollution flows can be seen in Figure 5 below.

Wonderland is quite smooth. Up to the mid-2070s, it appears to offer no particular difficulties for the forecaster. There are no baby booms and busts, no depressions or wars. There is just a smooth segment of an upward-sloping S-shaped curve. Adopting the idea in Rogers' (1995) of rating forecasting situations with respect to their degree of difficulty, we would rate the problem of predicting Wonderland's population with 95 percent confidence under the EDS as being comparatively simple. Wonderland's population history from 1930 to 1995 appears, for example, to be considerably more regular than the population history of the United States. The problem of forecasting Wonderland's population under the ENS after 2076, when it begins to experience a dreadful environmental crisis, is undoubtedly much more difficult. Below we will see how well various forecasting techniques perform on both problems.

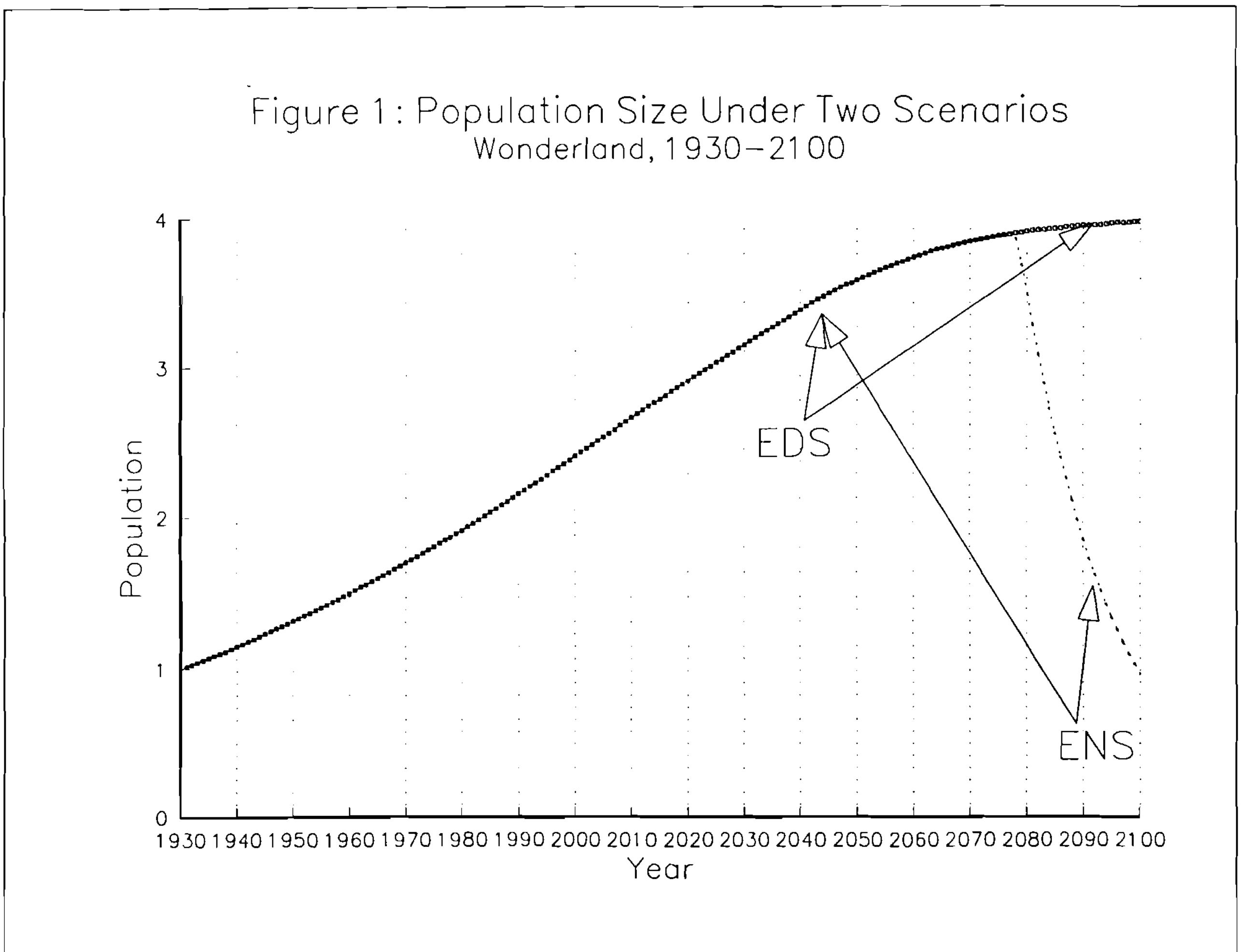

Figure 1. Population size under two scenarios. Wonderland, 1930-2100.

There are no structural changes in the Wonderland model around 2075. The path of Wonderland's population under the ENS follows from its history as continuously as it does in the case of EDS. In Wonderland, 147 years of common demographic history (from 1930 through 2076) does not imply even roughly similar demographic futures. In Wonderland this happens because population size is determined as part of a system that includes economic and environmental factors. From 1930 through the mid-2070s, the two scenarios are demographically, and economically identical, but are not identical with respect to their impacts on the environment. After 2076, the environmental impacts under ENS become significant for the first time, and then the demographic and economic paths of the two scenarios diverge.[4]

Figure 2 shows the time paths of the crude birth and death rates. These are the determinants of the population growth seen in Figure 1. The crude birth rates and death rates are assumed to be negatively related to the level of per capita income, and the crude death rate is also assumed to depend on the level of the stock of natural capital. Again, there are no baby booms or baby busts to predict here. Everything is reasonably smooth and linear from 1930 through 2076 under both scenarios. The behavior under the EDS after 2076 appears to be a simple extrapolation of what came earlier. The crisis under ENS, after 2076, is clear. The crude death rate skyrockets and even the crude birth rate increases, but not enough to keep the population size from decreasing.

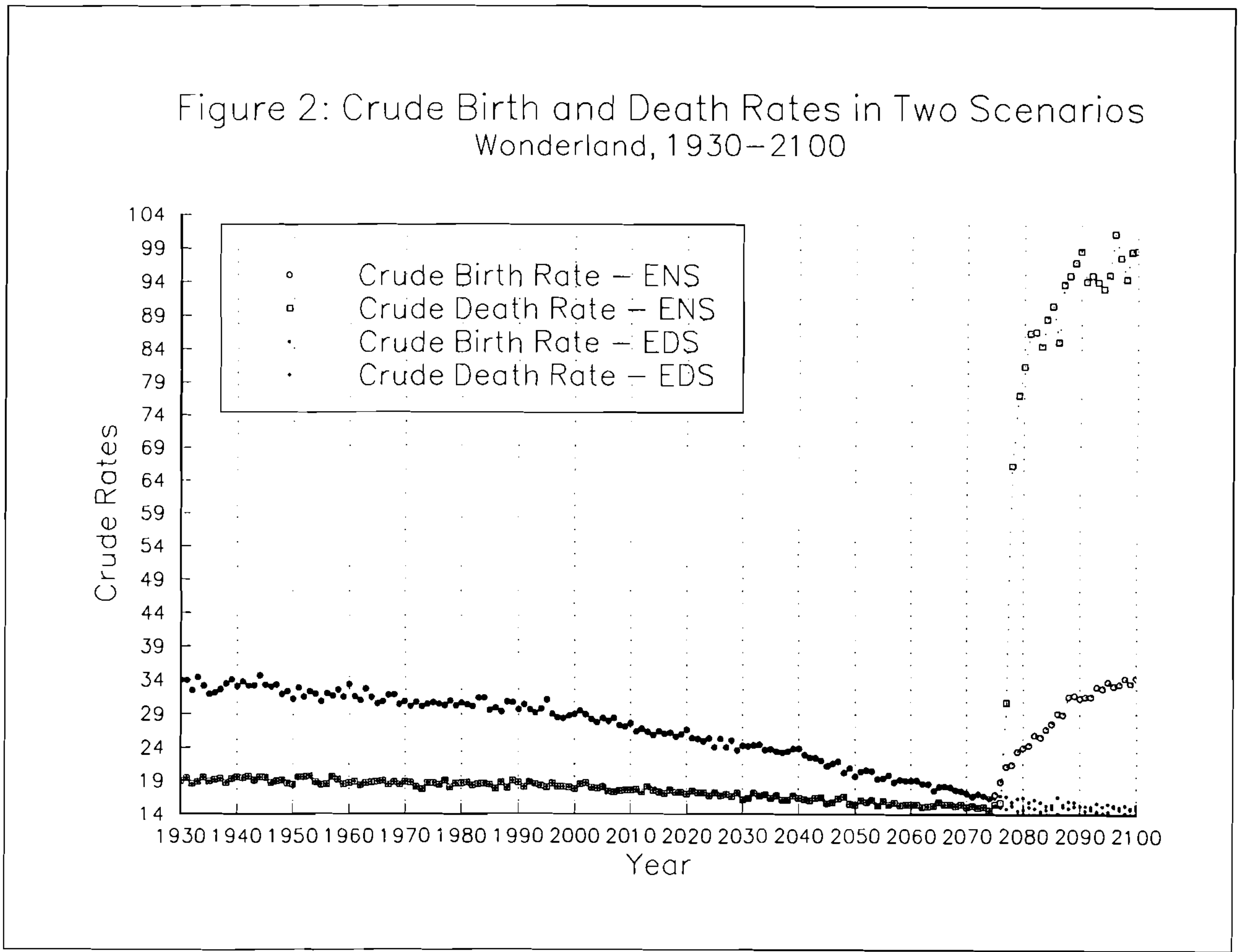

Figure 2. Crude birth and death rates in two scenarios. Wonderland, 1930-2100.

[4]Perhaps the idea of two future scenarios that are consistent with a single past can be made more concrete by thinking about the Earth. If we continue on our present course, there may be an environmental crisis around 2075 or there may not be. Today, we do not know which will occur. Currently, both scenarios would produce the same demographic and economic outcomes, because the environment is not polluted enough to differentiate the two.

Figure 3 shows per capita net output under the two scenarios. Per capita net output is total per capita output minus per capita pollution abatement costs. Per capita output has a constant rate of growth when Wonderland is unpolluted. When the stock of natural capital falls, economic growth slows down, and if it falls low enough, this decline in the rate of growth turns into an actual shrinkage in the economy.

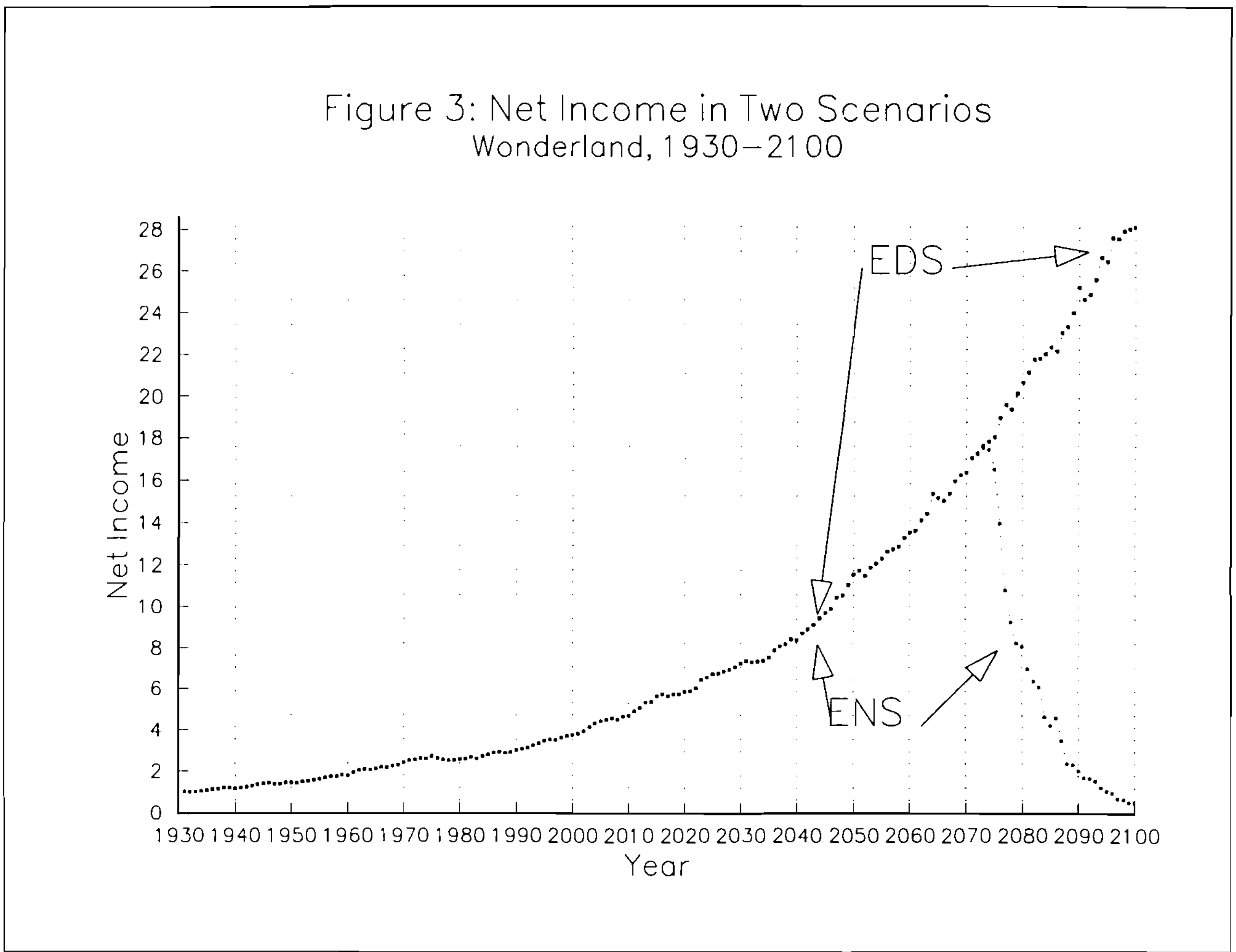

Figure 3. Net income in two scenarios. Wonderland, 1930-2100.

Figure 4 shows the stock of natural capital. The stock of natural capital is allowed to vary between a value of 1.0, in the case where the environment is sufficiently clean that it does not affect death rates or economic performance, to 0.0, in the case where a polluted environment has its worst possible effects on them. The environment has the ability to cleanse itself to some degree. This ability, however, is assumed to diminish as the amount of pollution increases.

From 1930 through the middle of the 2060s, the stock of natural capital remains at 1.0. In other words, before the mid-2060s, even though there is a great deal more pollution under ENS, there was not so much more that nature could not cleanse itself. It is only after the mid-2060s that the burden of pollution becomes too damaging under ENS, and the stock of natural capital falls precipitously. This leads to the massive increases in death rates and declines in net incomes that we saw previously. Under EDS, pollution remains under control and the stock of natural capital remains high throughout the period.

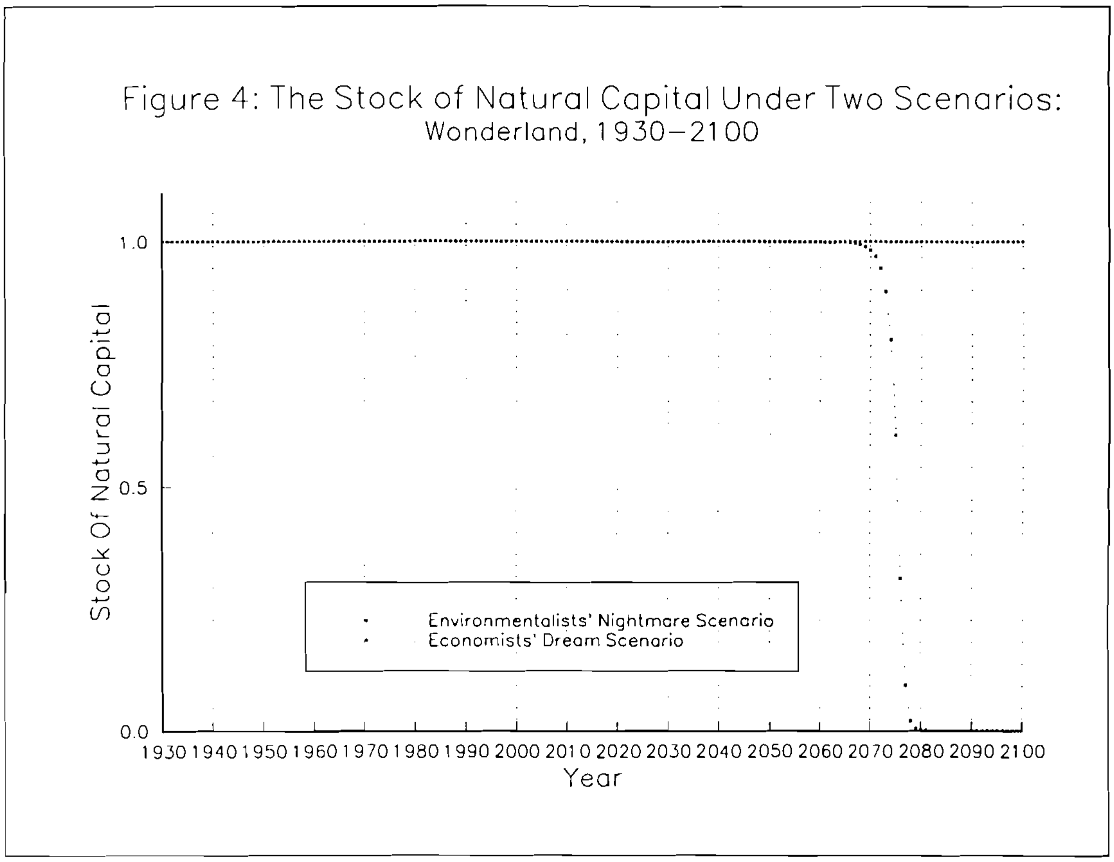

Figure 4. The stock of natural capital under two scenarios. Wonderland, 1930-2100.

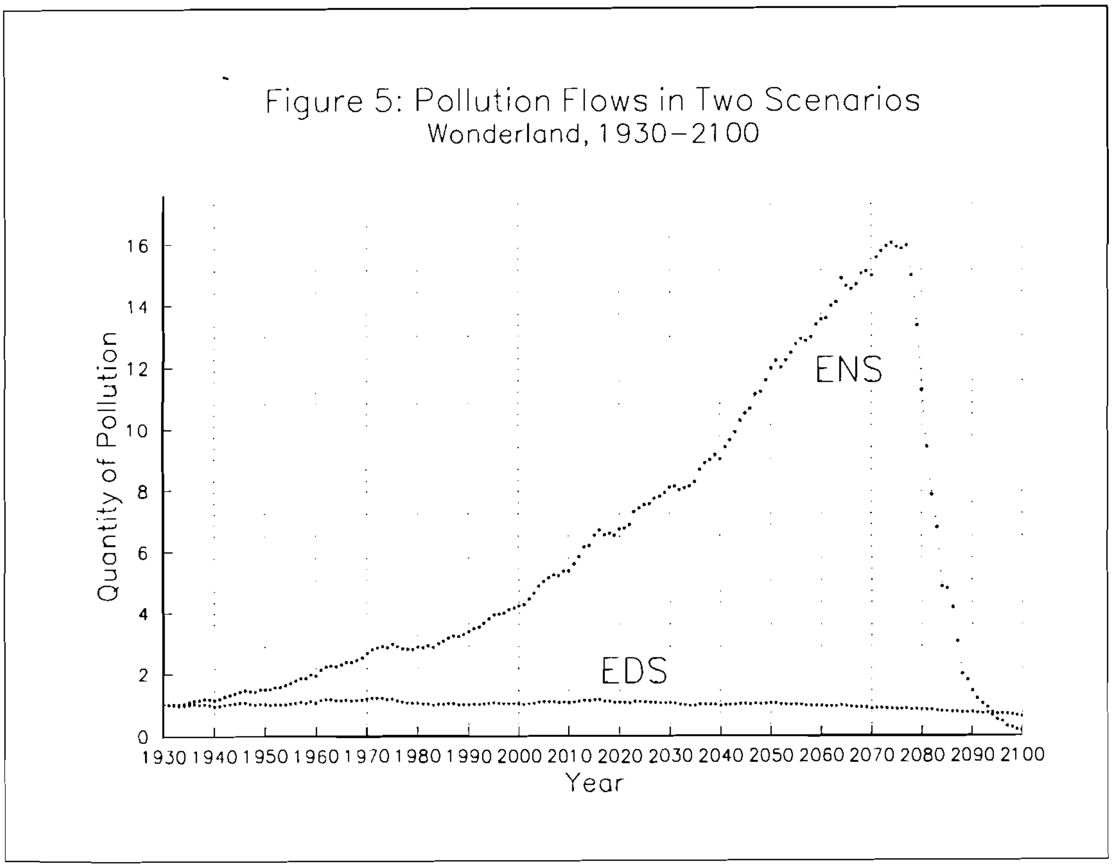

Figure 5. Pollution flows in two scenarios. Wonderland, 1930-2100.

Figure 5 shows the pollution flows on both sides of the looking-glass. We do not know what the rate of pollution per unit output will be in the future. We obtain the ENS by setting the parameter controlling that rate to a relatively high value. The implication of this possibility, as shown in the figure, is that pollution flows would rise continuously, so that by the beginning of the crisis they are around 16 times higher than they were in 1930. We obtain the EDS by setting that parameter to a relatively low level (see footnote 3 above for the exact parameter settings). The consequence of that assumption is that pollution goes up a little from its 1930 level and then goes down a little. It does not exhibit the substantial increase seen under ENS.

4. True 95 Percent Confidence Intervals

In order to produce true confidence intervals for Wonderland's future population, we first need a past. To obtain this past, we ran the Wonderland model once, from 1930 to 1995, given the model, parameters, and initial conditions in Appendix A. We used the ENS assumption in footnote 3 above, although the demographic histories from 1930 through 1995 are the same under both scenarios.[5]

Given the initial conditions in 1995 that are part of Wonderland's history,[6] we generated 3,200 stochastic futures. The true 95 percent confidence intervals shown in Table 1 are derived from those futures. The confidence intervals are expressed in terms of percentage growth from Wonderland's known 1995 population. We are quite certain about Wonderland's population size for the next 30 years. By 2025, we are 95 percent confident that Wonderland's population will grow by between 31.7 percent and 36.8 percent. For 2050, the confidence interval lies between 51.1 and 67.7 percent growth. This expected growth is the same on both sides of the looking-glass.

By 2075, the confidence intervals differ across scenarios, as crises are beginning to occur in the ENS. In the economists' dream scenario, we are 95 percent confident of population growth between 59.2 and 92.7 percent. In the environmentalists' nightmare scenario, we are 95 percent confident of population growth between -7.8 percent and 91.9 percent. The lower bound of the 95 percent confidence interval already shows a population decrease compared to 1995, indicating that crises must have begun quite a few years earlier. The upper bound, however, is quite close to that of the EDS. Even under the ENS, there are some futures in which serious crises do not materialize. In the economists' dream scenario, we are 95 percent confident that Wonderland's population will have grown between 60.8 and 105.3 percent by 2100. Under the ENS, we are exactly as confident that the population will have experienced something between a shrinkage of 80.7 percent and a growth of 103.2 percent. Of course, we are not at all confident which of those two confidence intervals is the correct one.

---

[5]All the data produced for this paper are available from the author upon request.

[6]The population as of January 1, 1995 was 2.2813. The rate of pollution for the year 1995 was 0.52, and income during 1995 was 3.4524.

Table 1. Complete and incomplete confidence intervals for Wonderland's population growth, 2025 to 2100.

Type of Prediction	Dependent Variable	Year			
		2025	2050	2075	2100
ENS	Population				
Upper		36.8	67.7	91.9	103.2
Average		34.2	59.3	67.0	-30.5
Lower		31.7	51.1	-7.8	-80.7
EDS	Population				
Upper		36.8	67.7	92.7	105.3
Average		34.2	59.3	75.2	80.7
Lower		31.7	51.1	59.2	60.8
Quadratic	Population				
Upper		38.0	74.9	115.9	159.4
Average		37.4	74.1	114.8	158.1
Lower		36.8	73.2	113.8	156.9
ARIMA (0,2,1)	ln(Population)				
Upper		56.5	139.2	273.3	499.1
Average[a]		35.5	67.9	101.1	133.7
Lower		17.3	17.9	8.3	-8.8
ARIMA (4,2,0)	ln(Population)				
Upper		46.0	106.8	197.0	339.4
Average[a]		39.5	84.0	142.8	220.3
Lower		32.8	63.7	98.4	138.8
ARIMA (0,1,1)	ln(CBR) ln(CDR)				
Upper		43.0	91.1	153.2	230.7
Average		39.4	80.5	130.2	189.1
Lower		35.7	70.5	110.2	153.1
Causal	CBR/CDR				
Upper		40.8	85.6	144.4	222.2
Average		39.6	83.5	141.3	217.4
Lower		38.4	81.6	138.2	212.9

Notes for Table 1:
Upper refers to the upper bound of the 95 percent confidence interval.
Average refers to the arithmetic average except in those rows marked with an [a]. In those cases, the average refers to the geometric average.
Lower refers to the lower bound of the 95 percent confidence interval.

ENS and EDS confidence intervals have been computed on the basis of 3,200 simulated futures. The first forecast year is 1996. ENS and EDS share the same history that was used in estimating the parameters for the other specifications.

All five specifications have been estimated on the basis of historical data from 1930 through 1995. Confidence intervals have been computed on the basis of 1,600 simulated futures.

More information about the specifications can be found in Appendix B.

5. Projections in Wonderland Based on Data From 1930 Through 1995

Now that we have the true confidence intervals for Wonderland's population, the task of this section is to ascertain how well the 95 percent confidence intervals derived from various stochastic specifications compare with the true confidence intervals. Since there will be a number of different confidence intervals appearing shortly, it is useful to make a distinction between the true confidence intervals generated by the Wonderland model and estimated confidence intervals. For this purpose, we will call the confidence intervals, based on the Wonderland model, complete confidence intervals, because they are based on a complete knowledge of the model used by nature to produce future populations. The complete confidence intervals are the true confidence intervals, a terminology that we will also use. Confidence intervals that are based on an incomplete model of reality will be called incomplete confidence intervals. When we speak of being completely confident or incompletely confident that the population will be within a specific range, we will be making the same distinction. Our task, expressed in this language, then, is to compare complete and incomplete confidence intervals.

In order to compare those confidence intervals, we have chosen five specifications. The first is what Rogers (1995) calls a simple extrapolative model.

$$N_t = \alpha_0 + \alpha_1 \cdot N_{t-1} + \alpha_2 \cdot (N_{t-1})^2 + \epsilon_t \qquad \textbf{(Spec. 1)}$$

recalling that N_t is just Wonderland's population in year t.

Here population size in a given period is just a quadratic function of its size in the previous period. Specification 1 was estimated by OLS over the 65 year period from 1931 through 1995. Given the estimated coefficients, the estimated variance of the error term, and the assumptions that the errors are normally distributed with a fixed variance and independent of one another, we produced population projections for the years 1996 through 2100. By repeating this process 1,600 times with different random draws, we produced the distribution of population projections for each year. The highest forty of those observations for each year (or 2.5 percent of the entire set of observations) lie above the upper bound of the 95 percent incomplete confidence interval, and the lowest forty observations lie below the lower bound of the 95 percent incomplete confidence interval. The regression coefficients and their standard errors appear in Appendix B, along with coefficients, standard errors and other pertinent information about the other specifications.

Specification 2 is slightly more complex. Here we used ARIMA methods for producing projections. First through fourth differences in the level of population showed signs of nonstationarity, either because of trends in the differenced variable or trends in its variance. The second difference of the logarithm of population appeared stationary. We used three criteria to choose the appropriate number of autoregressive and moving average terms, Akaike's original information criterion (Akaike 1973), Akaike's reformulated information criterion (Akaike 1981), and Sawa's BIC criterion (Sawa 1978). We systematically estimated the 15 model specifications with 3 or fewer autoregressive terms and three or fewer moving average terms. On all three criteria, the best choice was a model with a single moving average term:[7]

$$p_t = \alpha_0 + \alpha_1 \cdot \mu_{t-1} + \mu_t \qquad \textbf{(Spec. 2)}$$

where,
p_t is the second difference of the natural logarithm of population at time t, and
μ_t is a random error term at time t.

Specification 3 is just a different ARIMA model. Instead of choosing the model that was best on the three information criteria, we systematically searched over all models with 8 or fewer autoregressive terms (and no moving average terms). On the basis of the information criteria, the best in that group was not as good as the simple ARIMA(0,2,1) model. It was an ARIMA(4,2,0) model:

$$p_t = \alpha_0 + \alpha_1 \cdot p_{t-1} + \alpha_2 \cdot p_{t-2} + \alpha_3 \cdot p_{t-3} + \alpha_4 \cdot p_{t-4} + \mu_t \qquad \textbf{(Spec. 3)}$$

Specification 3 is included in order to see the effects of model choice in the ARIMA framework. In the case of both Specification 3 and 4, analytic (as opposed to simulated) confidence intervals are reported here.

Specification 4 is somewhat more complex. Here separate ARIMA models are used to estimate the crude birth and death rates, and those rates are used to produce stochastic population projections. In the cases of both the crude birth and death rates, we used the first difference of the logarithms.[8] The information criteria tests suggested again that the best model had a single moving average term and no autoregressive ones, so we used those.

$$\begin{aligned} b_t &= \alpha_0 + \alpha_1 \cdot \mu_{t-1} + \mu_t \\ d_t &= \beta_0 + \beta_1 \cdot \epsilon_{t-1} + \epsilon_t \end{aligned} \qquad \textbf{(Spec. 4)}$$

where,
b_t is the first difference of the natural logarithm of the crude birth rate in period t,
d_t is the first difference of the natural logarithm of the crude death rate in period t.

The equations were estimated separately, and in the projections, the error terms were assumed to be independent. The confidence intervals in Table 1 are computed from 1,600 stochastic futures.

---

[7]For all the ARIMA models tested here, the Box-Ljung Q-statistic was computed for 12, 18, 24, and 30 lags. In none of them could the null hypothesis of white noise residuals be rejected at the 10 percent level of significance, except in one case that will be noted below. Dickey-Fuller tests for all the ARIMA specifications rejected the hypothesis of a unit root at a 95 percent level of confidence.

[8]Logarithms were used to ensure against negative crude rates.

The final specification is a causal one; indeed, it is the true one in equations (A1), (A4) and (A5) in Appendix A, to the extent that it can be estimated from Wonderland's historical data.[9] In this framework, the crude birth and death rates are contemporaneously correlated because, in each year, they both depend on the same value of net income. The confidence intervals were again calculated from 1,600 simulated futures.

Using the Rogers' (1995) typology, the five specifications above range from a very simple extrapolative model to a more complex causal one. None of them are very complicated. Let us now turn to see how well these specifications do in predicting the true 95 percent confidence intervals for Wonderland's population.

Under either scenario, Wonderland's population will grow by between 31.7 and 36.8 percent by 2025, and between 51.1 and 67.7 percent by 2050. Thus, predicting population to 2050 is relatively easy. After 2050, when the confidence intervals diverge, the interpretation of the predictions becomes strained because although both scenarios are consistent with the same past demographic data, they offer distinctly different futures.

According to the quadratic specification, we are 95 percent incompletely confident that by 2025 Wonderland's population will grow by between 36.8 and 38.0 percent. Note that the lower bound of the 95 percent incomplete confidence interval is the upper bound of the complete interval. Indeed, beginning in 2025, the 95 percent incomplete confidence interval for the quadratic specification never even overlaps with the true confidence interval. Therefore, we are 95 percent confident that the population of Wonderland will lie OUTSIDE of the 95 percent incomplete confidence interval based on the quadratic specification.

The incomplete confidence intervals from the quadratic model are very narrow. In 2025, its width is only 1.2 percent of Wonderland's 1995 population. The width of the complete confidence interval is over four times larger. In 2050, the width of the incomplete confidence interval is about one-tenth that of the complete confidence interval. Using the quadratic specification, we are apparently even quite certain about what Wonderland's population would be in 2100. It tells us that we are 95 percent incompletely confident that the population would grow between 156.9 percent and 159.4 percent. Even under the EDS, Wonderland's population grows only between 60.8 percent and 105.3 percent by 2100. Use of the quadratic specification makes us feel highly confident of quite wrong figures.

The ARIMA(0,2,1) specification for the natural logarithm of population size has just the opposite characteristic. Its confidence intervals are much too wide. Even in 2025, the width of its confidence interval is already 39.2 percent of Wonderland's 1995 population. We are apparently 95 percent confident that Wonderland's population would grow between 17.3 percent and 56.5 percent. In truth, we are 95 percent confident that Wonderland's population will grow between 31.7 and 36.8 percent. In 2025, the width of the incomplete confidence interval is over 7 times that of the complete confidence interval. In 2100, the upper bound of the 95 percent incomplete confidence interval is 499.1 percent population growth. This is around 5 times higher than the true upper bounds in either the EDS or ENS. In all the specifications that we have

---

[9]Because the stock of natural capital is always equal to unity in the 1930-1995 period, pollution control costs are equal to zero (see equation (A8)), and there is no difference between per capita output and net per capita output. Therefore, equation (A1) can be used to calculate the net income that enters into the crude birth and death rate equations. For the same reason, Δ_3 and θ, the parameters in the death rate equation that are related to the stock of natural capital, cannot be computed. The functional forms that were estimated and the results appear in Appendix B.

estimated, including a number not reported here, we have observed a general tendency for ARIMA models to produce incomplete confidence intervals that are larger than their complete counterparts, sometimes quite a bit larger.

The ARIMA(4,2,0) specification for the natural logarithm of population has confidence intervals that are smaller than those in the previous ARIMA model, but still considerably larger than the true ones. In 2025, the 95 percent incomplete confidence interval lies between 32.8 and 46.0 percent population growth. The true interval is between 31.7 percent and 36.8 percent. The two lower bounds are close, but the incomplete upper bound is 9.2 percentage points too high. For 2050, the upper bound on the incomplete confidence interval is 106.8 percent growth, while the true upper limit is 67.7 percent growth. The average predictions of the (4,2,0) model are much worse, however, than those from the (0,2,1) model.

Specification 4 is the first of our two structural models. Among all our specifications, it is the one where the size of its 95 percent incomplete confidence intervals are closest to the true intervals. Unfortunately, the mean predictions are not very good. By 2050, for example, the mean predicted population growth is 80.5 percent, as compared to the true average of 59.3 percent. Worse, in a sense, is the fact that the incomplete 95 percent confidence intervals and the complete ones are disjoint after 2050. We are quite confident that the population of Wonderland in the second half of the 21st century will lie outside of the 95 percent confidence interval produced by Specification 4.

The final specification is the true structure, ignoring the effects of the environment. It may be surprising or disappointing to some to learn that the true structure does rather poorly. It has confidence intervals that are much too small, and it predicts average growth rates that are among the worst in the group. Why does the true structure do so poorly? The answer is that in the period 1930 to 1995, there is a great deal of noise in the crude birth and death rate series relative to the amount of variation. As a result, the parameters cannot be estimated very accurately even when we know the true structure. We are well over 95 percent confident that, for most of the 21st century, the true population of Wonderland would lie outside of the 95 percent incomplete confidence interval based on the causal model.

The population projections out to the year 2075 and beyond are highly suspect, both in the average predicted growth and in their estimated confidence intervals. Table 1 shows that estimated confidence intervals for long-run projections are themselves so uncertain as to provide us with no useful information.

If we were only to consider projections to the year 2025, we can see that the average predictions are reasonably close to the Wonderland's average outcome. The worst prediction, the one made by the model where the true structure was estimated, was for an average population growth of 39.6 percent, when the true average population growth is 34.2 percent. This is not bad. The confidence intervals are predicted much less accurately. In one specification, we are 95 percent confident that the population growth would be between 36.8 and 38.0 percent. In another, we are 95 percent confident that the population growth would be between 17.3 percent and 56.5 percent. Looking across specifications, it appears that we are much more uncertain about our confidence intervals than we are about the average amount of population growth itself. Confidence intervals are themselves estimates. Table 1 suggests the possibility that they may often be quite poor approximations to the true confidence intervals.

It is worth observing here that the great variability in incomplete confidence intervals in Table 1 is not due to differences in inputs. All the models are estimated using exactly the same data. Table 1 shows that, even in this case, different specifications can produce confidence intervals of very different size.

Do simpler models outperform more complex ones? This question is difficult to answer because performance here has two dimensions, the average forecast and its 95 percent incomplete confidence interval. The simplest model, the quadratic specification, gives the second closest prediction of expected population growth by 2050 for example, but its 95 percent incomplete confidence interval is too small by a factor of about 10. The most complex model is the causal specification. It has incomplete confidence intervals that are larger than those for the quadratic, but still too narrow. In addition, its predictions of expected population growth are among the poorest.

The ARIMA models stand intermediate in complexity. The most complex of these, the one based on estimated crude birth and death rates has reasonably sized incomplete confidence intervals, but predicted expected population growth worse than the quadratic. The ARIMA(0,2,1) model, the one with a single moving average term, predicted expected population growth best among our specifications, but it has the largest incomplete confidence intervals. In 2050, the incomplete confidence interval for that specification is over 7 times larger than the true interval.

On the basis of the models in Table 1, there is certainly no monotonic relation between simplicity and performance. Therefore, the choice between simpler or more complex models has to be made on other grounds.

6. Projections in Wonderland Based on Data From 1855 to 1995

Can we feel more secure in producing confidence intervals for shorter range forecasts? In order to investigate this, imagine that instead of starting our Wonderland database in 1930, we started it in 1855. We can now associate our 171 years of data with the years 1855 to 2025. We can produce 141 years of historical data (i.e. from 1855 through 1995), just as we produced the 66 years of history in Section 4, use those data to reestimate the specifications above and make projections out to the year 2025. Further, let us assume that we are living in the environmentalists' nightmare scenario. For reference, 1995 now is the point labelled 2070 in Figures 1 through 5. Clearly, we are on the verge of catastrophe, but we have had 141 years of continuous population growth, and continuous declines in the crude birth and death rates. This on the Rogers' (1995) scale would be a difficult test indeed.

On the basis of the 141 year history, we have estimated the parameters of four of the specifications that appear in Table 1, with the exception of the ARIMA(4,2,0) model. The figures in Table 2 are in terms of Wonderland's population size, not rates of growth as in Table 1. We think that it is easier to read this way. Given our redated database, Wonderland's 1995 population is 3.849. As can be seen from Table 2, in 10 years, that is, by 2005, Wonderland's population decreases, on average, to 3.336, or by 13 percent. Wonderland has begun a terrible environmental crisis. Death rates have begun to climb, and even with somewhat higher birth rates, the population is still shrinking. This dismal story continues to unfold decade by decade. Between 2005 and 2015, the population of Wonderland falls by 48 percent, and between 2015 and 2025 by another 48 percent. Moreover, we are not very uncertain about this catastrophe. In 2025, we are 95 percent completely confident that the population of Wonderland would be between 0.87 and 0.94.

Table 2. Complete and incomplete confidence intervals for Wonderland's population, 2005, 2015, and 2025.

Type of Prediction	Dependent Variable	Year		
		2005	2015	2025
ENS	Population			
Upper		3.458	1.825	0.944
Average		3.336	1.750	0.902
Lower		3.232	1.686	0.866
Quadratic	Population			
Upper		3.974	4.069	4.145
Average		3.962	4.050	4.128
Lower		3.950	4.039	4.111
ARIMA (0,2,1)	ln(Population)			
Upper		4.034	4.277	4.557
Average[a]		3.920	3.955	3.955
Lower		3.809	3.658	3.432
ARIMA CBR (0,1,1) CDR (1,1,2)	ln(CBR) ln(CDR)			
Upper		3.952	4.052	4.142
Average		3.920	3.969	3.998
Lower		3.889	3.891	3.865
Causal	CBR/CDR			
Upper		3.858	3.867	3.878
Average		3.847	3.852	3.860
Lower		3.837	3.835	3.840

Notes for Table 2:
Upper refers to the upper bound of the 95 percent confidence interval.
Average refers to the arithmetic average except in the row marked with an [a]. In that case, the average refers to the geometric average.
Lower refers to the lower bound of the 95 percent confidence interval.

ENS confidence intervals have been computed on the basis of 3,200 simulated futures. The first forecast year in 1996.

All four specifications have been estimated on the basis of historical data from 1855 through 1995. Confidence intervals have been computed on the basis of 1,600 simulated futures.

More information about the specifications can be found in Appendix C.

Now suppose population projections were made in 1995 on the basis of the past 141 years of data. According to the quadratic specification, we expect Wonderland's population to lie between 3.95 and 3.97 in the year 2005. Clearly, this range does not include the true 95 percent confidence interval. Within 10 years, the complete and the incomplete confidence intervals are disjoint. This specification predicts the highest population growth for Wonderland of those considered here and provides narrow confidence intervals around those wrong figures.

The ARIMA(0,2,1) model again produces extremely wide incomplete confidence intervals. It tells us that we are 95 percent incompletely confident that Wonderland's population in 2025 would be between 3.43 and 4.56, but in this case, even such a wide interval does not include Wonderland's true population. The ARIMA model based on predictions of crude birth and death rates produces smaller incomplete confidence intervals than the ARIMA(0,2,1) model with about the same expected populations.

The causal model of demographic processes, excluding the environment now does poorly for a completely different reason. The poor performance of causal model in Table 1 was due to the inaccuracy of the parameter estimates. Certainly, with 141 years of data, we should be able to estimate the parameters with much greater precision, and indeed this is the case. The estimates of parameters that relate net output to crude birth and death rates are now quite accurate, as is the estimate of the rate of growth of output in an unpolluted environment. The problem arises because the interactions between the environment, population, and the economy are ignored. The stock of natural capital has a value of 1.000 from 1855 through 1986, falls to 0.990 in 1993 and to 0.971 in 1995. The recent fluctuations seem to be small and it is difficult to know how to model them. Therefore, we followed a course that was very tempting, given the difficulty of incorporating environmental variables into our model; we ignored them. The result of ignoring environmental variables in all of the specifications here is to make predictions that immediately prove to be wrong.

Several things are evident from Table 2. First, the predictions and confidence intervals are extremely inaccurate, even 10 years into the future. It may be argued at this point that the models have been treated unfairly. Of course, a meteorite could strike the Earth next year, wiping out a quarter of its population. Forecasts are made assuming that everything stays the same in the future as it was in the past. This objection is wrong. In Wonderland, everything is exactly the same before 1995 and after 1995. No meteorite hits Wonderland, nor does anything else change exogenously. It is crucial to emphasize here that Wonderland is a continuous model. There are no step-functions and no thresholds. The problem with the objection is that it implicitly assumes that everything staying the same means that everything changes in a roughly linear manner. Even with a constant structure, there can be periods of more rapid and less rapid change. In nonlinear systems that are poorly understood, it is not uncommon to observe surprising periods, when the behavior of the system deviates from what one would have expected on the basis of linear extrapolations. This is not the same as an exogenous change, such as a collision with a large meteorite.

The lesson that can be derived from Table 2 is that forecasts are not necessarily good for any fixed number of years. A population forecast may be fine for 50 years, or it may be terrible in 5 years. There is no way that we can be truly 95 percent confident that the future will not hold any surprises for us in the next 25 or 50 years. Indeed, that itself would be very surprising.

7. Conclusions

In this paper, we have produced an artificial reality, called Wonderland, and have performed population projections on a dataset generated by it, where we know both the past and the future. We found that we are not at all confident of the 95 percent confidence intervals produced by standard statistical techniques. Those intervals are estimated with a great deal of error. Given the specifications that we used, it was often the case that the true population would, with 95 percent confidence, lie outside of the estimated confidence intervals. Some estimated confidence intervals were much too small, some much too big. There is no guarantee that we will be 95 percent confident that the true population will lie within the estimated 95 percent confidence intervals, even 10 years after the starting date. The longer the horizon of the projection the greater is the anticipated error in both the expected population size and its confidence intervals. There appears to be a tendency for confidence interval forecasts to degrade more rapidly than forecasts of average population size. Since we are not at all confident of our computed confidence intervals, it is not surprising that we often found that different specifications lead to disjoint 95 percent confidence intervals.

Given these conclusions, what should forecasters do? First, population forecasters should give up producing 95 percent confidence intervals around their estimates. These intervals are highly inaccurate and misleading. When a forecaster says that the 95 percent confidence interval around the forecast for some future date is between, say, 200 million people and 300 million, the user of that forecast interprets that statement to mean that it is quite likely that the population at that time would be between 200 and 300 million people. This is an incorrect interpretation. The truth of the matter is that the 200 and 300 million figures are themselves highly inaccurate. We saw in Tables 1 and 2 that we could truly be 95 percent confident that the population would be between 100 and 150 million, or between 240 and 260 million. Because the 95 percent confidence intervals are so poorly estimated, leading people to believe that future population sizes will lie within them is wrong.

If for some reason forecasters must produce 95 percent confidence intervals around their forecasts, another name for what they produce would be more appropriate. We favor calling those intervals "incomplete confidence intervals." This would at least warn users about how misleading those intervals could be. Still, there would seem to be very little use for incomplete confidence intervals. We would not necessarily prefer a model with smaller incomplete confidence intervals to a model with larger ones, because the smaller ones could be too small. Similarly, we would not necessarily prefer a model with broader incomplete confidence intervals because they may be too large. The size of incomplete confidence intervals provides no basis for evaluating models. It also tells us nothing about the uncertainty of the future.

Is it possible to obtain a measurement of the uncertainty of forecasts without 95 percent incomplete confidence intervals? Of course it is. The Keyfitz-Stoto[10] approach is still valid. We can estimate the extent of uncertainty in new forecasts on the basis of the performance of old ones. There are also variants of the Keyfitz-Stoto approach that can be used. If a new projection technique is invented, its uncertainty can be measured by using it to make projections of the present based on past data. The errors in these forecasts could provide us with a basis upon which to evaluate it. Finally, it is possible to have projection tournaments based on

---

[10]See Keyfitz (1981) and Stoto (1983).

simulated data.[11] None of this is as satisfying as knowing the true interval in which the future population will lie 95 percent of the time.

We grow wiser by learning more about the world. We also grow wiser by learning that there are some things that we cannot know. We cannot know the true 95 percent confidence intervals around our population projections. This may make us and the users of projections unhappy, but, unfortunately, there is nothing that we can do about it. We also grow wiser by accepting the inevitable.

References

Ahlburg, D.A. (1992) Error measures and choice of a forecast method: Lessons for forecasters. *International Journal of Forecasting* **8**:99-100.

Ahlburg, D.A., and Land, K.C. (1992) Population forecasting: Guest editors' introduction. *International Journal of Forecasting* **8**:289-300.

Akaike, H. (1973) Information theory and the extension of the maximum likelihood principle. Pages 267-281 in B.N. Petrov and F. Csaki, eds. *2nd International Symposium of Information Theory.* Budapest: Akailseoniai-Kiudo.

Akaike, H. (1981) Likelihood of a model and information criteria. *Journal of Econometrics* **16**: 3-14.

Alho, J.M. (1990) Stochastic methods in population forecasting. *International Journal of Forecasting* **6**:521-530.

Carter, L.R., and Lee, R.D. (1986) Joint forecasts of U.S. marital fertility, nuptiality, births and marriages using time series models. *Journal of the American Statistical Association* **81(396)**: 902-911.

Carter, L.R., and Lee, R.D. (1992) Modeling and forecasting U.S. sex differentials in mortality. *International Journal of Forecasting* **8**:393-411.

Cohen, J.E. (1986) Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches. *Demography* **23(1)**:105-126.

Keyfitz, N. (1981) The limits of population forecasting. *Population and Development Review* **7**: 579-593.

Lee, R.D. (1992) Stochastic demographic forecasting. *International Journal of Forecasting* **8**:315-327.

Lee, R.D. (1993) Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting* **9**:187-202.

Lee, R.D., and Carter, L.R. (1992) Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* **47(14)**:659-671.

[11]Such tournaments are discussed in Ahlburg and Land (1992) and suggested by Rogers (1995).

McDonald, J. (1979) A time series approach to forecasting Australian total live births. *Demography* **16(4)**:575-601.

McNown, R., and Rogers, A. (1989) Forecasting mortality: A parameterized time series approach. *Demography* **26(4)**:645-660.

McNown, R., and Rogers, A. (1992) Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting* **8(3)**:413-432.

McNown, R., Rogers, Andrei, and Little, Jani. (1995) Simplicity and complexity in extrapolative population forecasting models. *Mathematical Population Studies*. Forthcoming.

Mittnik, Stefan (1994) *Broccoli: Time Series the Healthy Way*. Version 1.1. Department of Economics. State University of New York, Stony Brook. Manuscript and computer program.

Pflaumer, P. (1988) Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting* **4(1)**:135-142.

Pflaumer, P. (1992) Forecasting US population totals with the Box-Jenkins approach. *International Journal of Forecasting* **8**:329-338.

Rogers, Andrei (1995) Population forecasting: Do simple models outperform complex models? *Mathematical Population Studies*. Forthcoming.

Saboia, J.L.M. (1974) Modeling and forecasting populations by time series: The Swedish case. *Demography* **11**:483-492.

Saboia, J.L.M. (1977) Autoregressive integrated moving average (ARIMA) models for birth forecasting. *Journal of the American Statistical Association* **72**:264-270.

Sanderson, Warren C. (1994) Simulation models of demographic, economic, and environmental Interactions. In Wolfgang Lutz, ed. *Population-Development-Environment: Understanding Their Interactions in Mauritius.* Berlin: Springer Verlag. Forthcoming.

Sawa, T. (1978) Information criteria for discriminating among alternative regression models. *Econometrica* **46**:1273-1291.

Stoto, M. (1983) The accuracy of population projections. *Journal of the American Statistical Association* **78**:13-20.

Appendix A: The Wonderland Model: Equations, Parameters, Initial Conditions, and Standard Deviations

A1. Equations

This description of the Wonderland model draws heavily on Sanderson (1994). Equations (A1) and (A2) deal with the economy, equations (A3) through (A5) with the population, equations (A6) and (A7) with the environment, and equations (A8) and (A9) with environmental policy. The subscript *t* in the equations refers to the year in which the variables are measured, all Greek letters refer to parameters of the model, all subscripted *u* terms refer to independently, normally distributed random variables, and v_t refers to a uniformly distributed random variable.

Notation:

I_t	per capita output in year t,
NK_t	stock of natural capital in year t,
NI_t	net per capita output in year t,
CBR_t	the crude birth rate in year t,
CDR_t	the crude death rate in year t,
N_t	population in year t,
PF_t	pollution flow in year t,
T_t	pollution per unit of output (the rate of pollution) in year t,
PC_t	per capita pollution control expenditures in year t.

Wonderland Model Notation

Equation (A1) says that per capita output in year t+1, I_{t+1}, depends on the level of per capita output in year t, I_t, the level of the stock of natural capital in year t, NK_t, two stochastic terms γ_t, η_t, and a parameter λ. Natural capital may be thought of as the set of things provided to us by the environment, like air and water, which allow us to live healthy and productive lives. When natural capital is undiminished by pollution, the stock is given a value of 1.0.[12] As more pollution occurs, the level of the stock of natural capital diminishes, until, in the extreme, it reaches 0.0. Equation (A1) incorporates the assumption that the lower the stock of natural capital, the lower the rate of per capita output growth. When NK_t is equal to 1.0, the rate of economic growth is γ_t; when NK_t is equal to 0.0, the rate of economic shrinkage is equal to η_t.

Equation (A2) simply defines net per capita output as the difference between the level of per capita output in equation (A1) and per capita expenditure on pollution control. The latter is determined by pollution control policy, as expressed by equation (A8). Standard economic accounting procedures do not put a (negative) value on pollution, but do include pollution control expenditures. Therefore, offsetting increases in pollution and pollution control expenditures, which leave stocks of natural capital unchanged, show up as increases in per capita output. To avoid this possibility, net per capita output is used in Wonderland as the measure of economic well-being.

[12]Equation A7 is undefined when *NK* is exactly 1.0. In all the runs of the Wonderland model, *NK* is assumed to be bounded from above by 0.99999.

Equations (A3), (A4), and (A5) make up the demographic portion of the model. Equation (A3) determines the crude birth rate (the ratio of births to the population multiplied by 1,000) and equation (A4) the crude death rate (the ratio of deaths to the population multiplied by 1,000). The model assumes that both crude rates decrease with increases in net per capita output. In addition, equation (A4) stipulates that decreases in the stock of natural capital cause the crude death rate to rise. Both equations (3) and (4) include stochastic terms. Equation (A5) is an accounting identity that computes the population in year $t+1$ based on the population in year t and the intervening numbers of births and deaths.[13]

The environmental portion of the model appears in equations (A6) and (A7). Equation (A6) determines the annual flow of pollutants as a function of the population size, per capita output, the technologies of production and pollution abatement, and on the amount of money that is spent on pollution control. Equation (A7) shows how the combination of the pollution flow and nature's ability to cleanse itself interact to produce changes in the stock of natural capital. Our specification is based on the idea that nature has the ability to cleanse itself, but that the strength of this ability decreases as the quantity of pollution increases.

The final equations represent environmental policy. In equation (A8), the stock of natural capital is constantly monitored and the amount of money per person spent on pollution control determined depending on per capita output, the stock of natural capital, and a random term. Per capita spending on pollution control is assumed to increase with per capita output and to increase as the stock of natural capital decreases. In equation (A9), the rate of pollution (pollution per unit of output) is assumed to be a random walk with a drift of χ percent per year.

[13]The model assumes that there is no net migration.

The Wonderland model:

Economy

$$I_{t+1} = I_t\left[1+\gamma_t-(\gamma_t+\eta_t)\cdot(1 - NK_t)^{\lambda}\right] \quad \text{(A1)}$$
where
$$\gamma_t = \gamma + u_{\gamma,t}$$
$$\eta_t = \eta + u_{\eta,t}$$

$$NI_t = I_t - PC_t \quad \text{(A2)}$$

Population

$$CBR_t = \mathrm{B}_1\cdot\left[\mathrm{B}_2 - \left(\frac{e^{\beta\cdot NI_t}}{1 + e^{\beta\cdot NI_t}}\right)\right]\cdot(1 + u_{cbr,t}) \quad \text{(A3)}$$

$$CDR_t = \Delta_1\cdot\left[\Delta_2-\left(\frac{e^{\alpha\cdot NI_t}}{1+e^{\alpha\cdot NI_t}}\right)\right]\cdot\left[1 + \Delta_3\cdot(1 - NK_t)^{\theta}\right]\cdot(1 + u_{cdr,t}) \quad \text{(A4)}$$

$$N_{t+1} = N_t\cdot\left[1 + \left(\frac{CBR_t - CDR_t}{1000}\right)\right] \quad \text{(A5)}$$

Environment

$$PF_t = N_t\cdot I_t\cdot T_t - \kappa\cdot\left(\frac{e^{\epsilon\cdot PC_t\cdot N_t}}{1+e^{\epsilon\cdot PC_t\cdot N_t}}\right) \quad \text{(A6)}$$

$$NK_{t+1} = \frac{e^{\ln\left(\frac{NK_t}{1-NK_t}\right) + \delta\cdot(NK_t)^{\rho} - \omega\cdot PF_t}}{1 + e^{\ln\left(\frac{NK_t}{1-NK_t}\right) + \delta\cdot(NK_t)^{\rho} - \omega\cdot PF_t}} \quad \text{(A7)}$$

Environmental Policy

$$PC_t = \phi\cdot(1 - NK_t)^{\mu}\cdot I_t\cdot(1 + u_{pc,t}) \quad \text{(A8)}$$

$$T_t = T_{t-1}\cdot(1 + \chi) + v_t \quad \text{(A9)}$$

A2. Parameters

These parameters differ somewhat from the parameters in Sanderson (1994). The only change that alters the behavior of the model is the change in ρ. The new parameter value keeps the model from cycling in the long-run.

Economy

$\gamma = 0.02$ $\eta = 0.10$ $\lambda = 2.0$

Population

$B_1 = 40$ $B_2 = 1.375$ $\beta = 0.16$
$\Delta_1 = 10$ $\Delta_2 = 2.5$ $\Delta_3 = 4$ $\alpha = 0.18$ $\theta = 15$

Environment

$\kappa = 2$ $\epsilon = 0.02$ $\delta = 1.0$ $\rho = 2.00$ $\omega = 0.10$

Environmental Policy

$\phi = 0.50$ $\mu = 2.0$

Technology

$\chi = -0.01$
Environmentalists' Nightmare Scenario

or

$\chi = -0.03$
Economists' Dream Scenario

A3. Initial Conditions

I = 1 NK = 1 N = 1 T = 1

In Sections 3, 4, and 5, these initial conditions are associated with the year 1930. In Section 6, these initial conditions are used for the year 1855.

A4. Standard Deviations

The standard deviations used are:

$u_{\gamma,t} = 1.00$ $u_{\eta,t} = 1.00$ $u_{cbr,t} = 0.02$ $u_{cdr,t} = 0.02$ $u_{pc,t} = 0.05$

v_t is a uniformly distributed random variable that ranges between -0.0025 and 0.0025.

Appendix B: Parameter Estimates of the Specifications in Table 1

Specification 1 - Quadratic
Dependent Variable - population size
See equation (Spec. 1) in the text

Parameter	α_0	α_1	α_2
Estimate	0.003539	1.012	-0.0009647
Standard Error	0.002938	0.004	0.001174

Estimated Standard Error of the Residual: 0.001149
Estimated by OLS Program: GAUSS

Specification 2 - ARIMA (0,2,1)
Dependent Variable - second difference of the natural logarithm of population size
See equation (Spec. 2) in text

Parameter	α_0	α_1
Estimate	-0.0000556	-0.9857
Standard Error	not avail.	0.021945

Estimated Standard Error of the Residual: 0.000760
Estimated by Maximum Likelihood Program: BROCCOLI

Specification 3 - ARIMA (4,2,0)
Dependent Variable - second difference of the natural logarithm of population size
See equation (Spec. 3) in text

Parameter	α_0	α_1	α_2	α_3	α_4
Estimate	-0.0001778	-0.8160	-0.6226	-0.4195	-0.3394
Standard Error	not avail.	0.1171	0.1467	0.1477	0.1212

Estimated Standard Error of the Residual:
Estimated by Maximum Likelihood Program: BROCCOLI

Specification 4 - ARIMA (0,1,1) for crude birth rate and crude death rate
95 Percent Confidence Intervals for Specification 4 Estimated Using GAUSS

Crude Birth Rate
Dependent Variable - first difference of the natural logarithm of the crude birth rate
See equation (Spec. 4) in text

Parameter	α_0	α_1
Estimate	-0.001404	-0.8570
Standard Error	not avail.	0.0665

Estimated Standard Error of the Residual: 0.02344
Estimated by Maximum Likelihood Program: BROCCOLI

Crude Death Rate

Dependent Variable - first difference of the natural logarithm of the crude death rate
See equation (Spec. 4) in text

Parameter	β_0	β_1
Estimate	-0.0007431	-0.8919
Standard Error	not avail.	0.0579

Estimated Standard Error of the Residual: 0.02098
Estimated by Maximum Likelihood Program: BROCCOLI

Specification 5 - Causal Model

Income

$$I_{t+1} = I_t[1+\gamma_t] \quad \text{where} \quad \gamma_t = \gamma + u_{\gamma,t} \tag{1A}$$

Parameter	γ
Estimate	0.01952
Standard Error	0.00292
True Value	0.02000

Estimated Standard Error of the Residual: 1.206
True Standard Error of the Residual: 1.000
Estimated by Nonlinear Least Squares Program: GAUSS

Crude Birth Rate

$$CBR_t = B_1 \cdot \left[B_2 - \left(\frac{e^{\beta \cdot NI_t}}{1 + e^{\beta \cdot NI_t}}\right)\right] \cdot (1 + u_{cbr,t}) \tag{2A}$$

Parameter	B_1	B_2	β
Estimate	15.518	2.8719	0.7783
Standard Error	1.1355	0.1741	0.2563
True Value	40.000	1.3750	0.1600

Estimated Standard Error of the Residual: 0.02207
True Standard Error of the Residual: 0.02000
Estimated by Nonlinear Least Squares Program: GAUSS

Crude Death Rate

$$CDR_t = \Delta_1 \cdot \left[\Delta_2 - \left(\frac{e^{\alpha \cdot NI_t}}{1+e^{\alpha \cdot NI_t}}\right)\right] \cdot (1 + u_{cdr,t}) \tag{3A}$$

Parameter	Δ_1	Δ_2	α
Estimate	4.4381	5.0620	0.5534
Standard Error	2.0355	2.1658	0.5676
True Value	10.000	2.5000	0.1800

Estimated by Nonlinear Least Squares Program: GAUSS

Appendix C: Parameter Estimates of the Specifications in Table 2

Specification 1 - Quadratic
Dependent Variable - population size
See equation (Spec. 1) in the text

Parameter	α_0	α_1	α_2
Estimate	-0.01320	1.032	-0.006626
Standard Error	0.001470	0.001	0.0002690

Estimated Standard Error of the Residual: 0.002069
Estimated by OLS Program: GAUSS

Specification 2 - ARIMA (0,2,1)
Dependent Variable - second difference of the natural logarithm of population size
See equation (Spec. 2) in text

Parameter	α_0	α_1
Estimate	-0.000092	-0.7889
Standard Error	not avail.	0.0531

Estimated Standard Error of the Residual:
Estimated by Maximum Likelihood Program: BROCCOLI

Specification 3 - ARIMA (0,1,1) for crude birth rate and ARIMA (1,1,2) for crude death rate

Crude Birth Rate
Dependent Variable - first difference of the natural logarithm of the crude birth rate
See equation (Spec. 4) in text

Parameter	α_0	α_1
Estimate	-0.004853	-0.6094
Standard Error	not avail.	0.0679

Estimated Standard Error of the Residual: 0.02393
Estimated by Maximum Likelihood Program: BROCCOLI

Crude Death Rate
Dependent Variable - first difference of the natural logarithm of the crude death rate

$$d_t = \beta_0 = \beta_1 \cdot d_{t-1} + \beta_2 \cdot e_{t-1} + \beta_3 \cdot e_{t-2} + e_t$$

Note: This equation differs from (Spec. 4) in text.

Parameter	β_0	β_1	β_2	β_3
Estimate	-0.0004672	0.7164	-1.7282	0.7815
Standard Error	not avail.	0.0883	0.07801	0.0684

Estimated Standard Error of the Residual: 0.01997
Estimated by Maximum Likelihood Program: BROCCOLI

95 Percent Confidence Intervals Estimated Using GAUSS

Specification 4 - Causal Model

Income

$$I_{t+1} = I_t[1+\gamma_t] \quad \text{where} \quad \gamma_t = \gamma + u_{\gamma,t} \tag{1A}$$

Parameter	γ
Estimate	0.02037
Standard Error	0.00176
True Value	0.02000

Estimated Standard Error of the Residual: 1.024
True Standard Error of the Residual: 1.000
Estimated by Nonlinear Least Squares Program: GAUSS

Crude Birth Rate

$$CBR_t = B_1 \cdot \left[B_2 - \left(\frac{e^{\beta \cdot NI_t}}{1 + e^{\beta \cdot NI_t}}\right)\right] \cdot (1 + u_{cbr,t}) \tag{2A}$$

Parameter	B_1	B_2	β
Estimate	39.548	1.3888	0.1653
Standard Error	0.5771	0.0141	0.0055
True Value	40.000	1.3750	0.1600

Estimated Standard Error of the Residual: 0.02028
True Standard Error of the Residual: 0.02000
Estimated by Nonlinear Least Squares Program: GAUSS

Crude Death Rate

$$CDR_t = \Delta_1 \cdot \left[\Delta_2 - \left(\frac{e^{\alpha \cdot NI_t}}{1+e^{\alpha \cdot NI_t}}\right)\right] \cdot (1 + u_{cdr,t}) \tag{3A}$$

Parameter	Δ_1	Δ_2	α
Estimate	10.000	2.5052	0.1897
Standard Error	0.3043	0.0622	0.0136
True Value	10.000	2.5000	0.1800

Estimated Standard Error of the Residual: 0.01850
True Standard Estimate of the Residual: 0.02000
Estimated by Nonlinear Least Squares Program: GAUSS