Veteran Population Model

Overview

Introduction

The Veteran Population Model (VetPop2007) – using combined data from the Department of Veterans Affairs (VA), Department of Defense (DoD), and Bureau of the Census – provides the official estimates and projections of the veteran population for each year from April 1, 2000, through September 30, 2036. This information is grouped into the following main categories: State, Age, Gender, Period of Service, Race/Ethnicity, Rank (Officer/Enlisted), and Branch of Service. Based primarily on a series of statistical analysis software (SAS) programs and data sets, the model stores both parameters and output tables in a format that users can manipulate easily via Microsoft Excel.

The Overview section is a part of the VetPop2007 Technical Documentation which covers a range of information about the model, from general description to technical details. Other sections document methodologies, pertinent assumptions, and data sources.

History

VetPop2007 represents the fifth generation of VetPop models:

- The original VetPop model, developed in the early 1990s before the Office of the Actuary (OACT) was established, relied heavily on data from the 1990 Census. Data from DoD's Defense Manpower Data Center (DMDC) provided information on monthly separations by Age, Gender, and State. At that time, the model ran on a mainframe computer at VA's Austin Automation Center.

- OACT's first population model was VetPop2000. Continuous enhancements were made in developing OACT's subsequent models: VetPop2001, completed in November 2001; VetPop2001Adj, completed in early 2003; VetPop2004, completed in December 2004; and VetPop2007, completed in December 2007.

- The VetPop2000 and VetPop2001 models began relying more heavily on DMDC data to project the population of veterans in the post-Vietnam era. The Census data were used to determine the number of Vietnam-era veterans (those who were age 39 or older on April 1, 1990, or who left service before May 8, 1975); the DMDC data were used to determine the number of post-Vietnam veterans (those who were age 38 or younger on April 1, 1990, or who left service after May 8, 1975).

- VetPop2001Adj (the official name is "VetPop2001 Adjusted to Census 2000") was a modification of the VetPop2001 model, with population totals adjusted to be consistent with control totals obtained from early releases of Census 2000 data (Summary File 3).

- VetPop2004 used more detailed Census 2000 data to provide more comprehensive and current estimates and projections.

What's New

VetPop2007 is consistent with the previous model primarily in that the estimate as of April 1, 2000, is based solely on Census 2000 veteran counts (custom tabulations) by single year of age, gender, state, and other characteristics. VetPop2007 reflects an update of several data sources and a few methodological enhancements.

- The model incorporates more recent DoD data on actual separations through September 30, 2006.
- VetPop2007 introduces a new summary table for Gulf War veterans, allowing the distinction of those who served since October 1, 2001.
- State to state migration incorporates American Community Survey (ACS) data to supplement the migration data from Census 2000. This improves state estimates by using more current information of veterans' whereabouts, as the accuracy of the Census 2000 data diminishes with time.
- Race estimates for the separations between 2000 and 2006 are based on DoD administrative files, rather than relying on the Census 2000 race distribution.
- VetPop2007 employs improved mortality by integrating additional data sources.

Input Data

Census 2000 and ACS

VetPop2007 retained the use of Census 2000 data as in VetPopo2004 to set the baseline estimate at April 1, 2000. During the time when VetPop2004 was under development (2003–2004), detailed Census 2000 data on veterans became available. The Census 2000 veteran data, referred to as "celled" data, contains special tabulations of the veteran population by Age, Gender, Period of Service, and Race/ethnicity at the State level. Because the data cells were rounded by the Census, OACT developed an algorithm to "de-round" the cells to make them match the published Census Summary File 4 (SF4) data. (You can find more detailed information on the special tabulations request and the de-rounding method used to remove the disclosure effect in the Appendix.)

DMDC

DMDC data, contained in two separate DoD files, are used to estimate the number of veterans separating after April 1, 2000, and through September 30, 2006. DMDC provided person level information on separations from the regular military in the Active Duty Loss Data File (often abbreviated as "Adloss" or referred to as the "Active Duty File"); separations from the Reserve Forces were extracted from the Reserve Components Common Personnel Data System file (referred to as the "Reserve File").

The Active Duty File maintains demographic and military data on all service members from the regular military who have separated from active duty since July 1, 1970. Coast Guard data is available beginning July 1, 1988. In recent years, the National Oceanic and Atmospheric Administration (NOAA) Corps and the Public Heath Service (PHS) also became part of the file. The data on the service members are fixed as of the day of separation.

The Reserve File contains personnel information on all present and past members of the Guard and Reserve Forces. According to VA, veterans are those who have been discharged from federal active duty. This excludes Reserve Forces whose only active duty was for training – unless disabled during training – and National Guard members whose only active duty was in state service. (You can find more detailed information on the Active Duty and Reserve components in the Data Sources section.)

GORGO

To estimate the number of veterans separating after September 30, 2006, VetPop2007 uses GORGO data. GORGO (not an acronym!) is the name of the projection model used by DoD's Office of the Actuary. This model annually projects active and separated military personnel to value retirement liabilities. GORGO also projects the military separations over a 100-year period. (You can find more detailed information in the GORGO part of the Data Sources section.)

Data summary

No single source can provide all the data needed to estimate the number of veterans separating from active service at different time periods. Figures 1 and 2 below illustrate this: first, by giving the range of years from which separation data are available; second, by showing the range of years actually used in the model

OVERVIEW I FIGURE 1. USE CENSUS DATA FROM 1975 – 2000 (VETPOP2007)
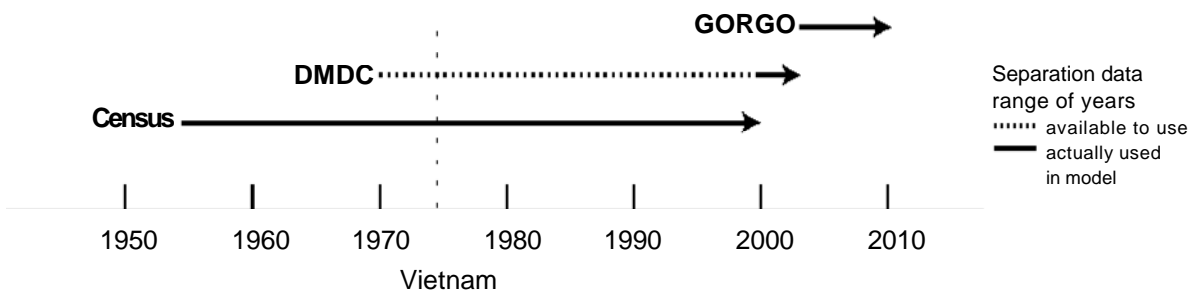
OVERVIEW I FIGURE 2. USE DMDC DATA FROM 1975 – 2000 (VETPOP2001ADJ)
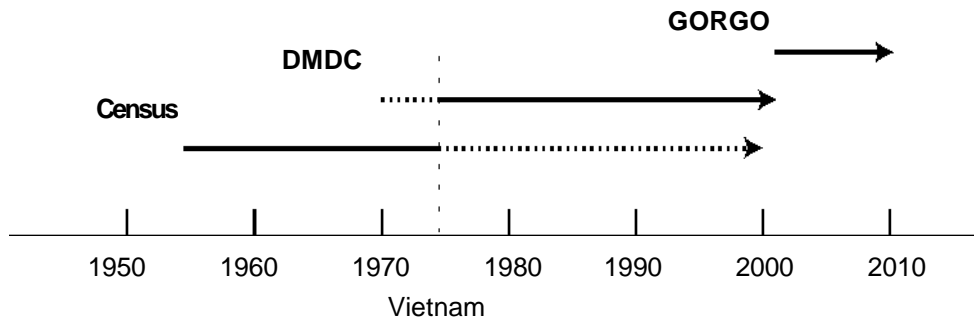
Parameter Files

Since none of the data sources in Figures 1 and 2 provide complete information needed by the model, VetPop2007 employs parameters to fill in, where needed, and to assist in integrating the data. These parameters are sets of demographic and actuarial assumptions about the veteran population. Because they are stored separately as Excel files that users can modify easily, OACT has the ability to rerun the model using new assumptions. The parameter files, listed in Overview Table 1., are described in detail in the Parameters section.

OVERVIEW I TABLE 1. VETPOP2007 PARAMETER FILES

Phase	Parameter Name	Purpose
Preprocessing	Branch_Census	Determine Branch for Census vets
	Mort_Census	Determine Deaths for Census vets (after 4/1/2000)
	Officer_Census	Determine Officer for Census vets
	Vets_Census	Adjust number of Veterans for foreign countries
Processing	Migrate_Historical	Migrate veterans for period 4/1/2000 ~~– 9/30/2006
	State_Seps	State distribution of new separations post 4/1/2000
	Vets_Dmdc	Adjust number of DMDC vets for veterans in Non-defense Branch
Projection	Branch_DoD	Determine Branch for future separations post 9/30/2006
	Gender_DoD	Determine Gender for future separations post 9/30/2006
	Mort (Mort_Comp)	Determine Deaths for projected vets
	Mort_Adjust	Adjust future mortality rates
	Migrate_Future	Migrate veterans for post 9/30/2006
	Race_DoD	Determine Race for new separations post 9/30/2006
	State_DoD	State distribution of new separations post 9/30/2006
	Vets_DoD	Estimate veterans who enter and leave service in the same fiscal year

Process

The VetPop2007 model can be broken up into three programming phases: preprocessing, processing, and projection. The detailed flowchart is included in the Appendix.

Preprocessing

The preprocessing phase pulls in the initial data sets, creates standardized variables, un-duplicates within DMDC files (to account for persons with multiple separations), and un-duplicates between Active Duty and Reserve Files. This last case accounts for overlap between two files, which may occur when an individual begins a career with the regular military and later transitions into the Reserve Forces.

Processing

In the processing phase, the first step is to generate the April 1, 2000, baseline, as indicated in the following table. The first column lists the fundamental (or "classification") variables. The next two columns represent Census 2000 data, which is used to count veterans who separated prior to April 1, 2000.

The letter "A" indicates that actual data is available, and the letter "P" indicates that a parameter file is used to "fill in" for missing variables. The data may be fully adjusted **(red)** or partially adjusted (blue and enclosed in parenthesis). You can refer to the Parameter section for more detailed information.
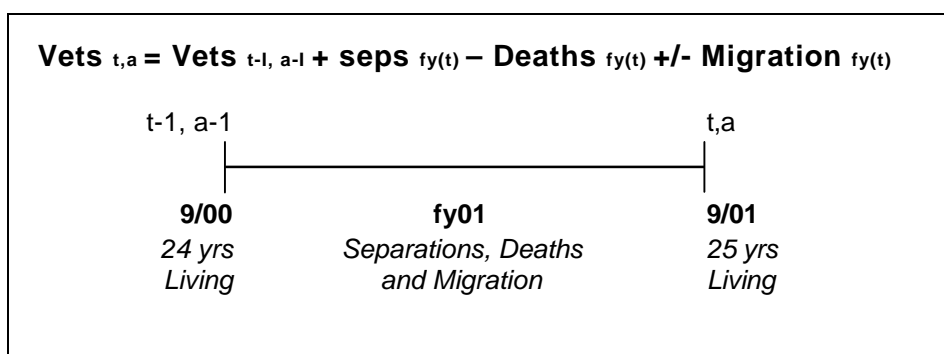
BUILDING THE APRIL 2000 BASELINE

Category	Census Parameter	Census Data Adjustment
Date	Pre-2000	
Age	A	
Gender	A	
Period of Service	A	
Race	A	
Veterans	A	(Vets_census)
Branch	P	**Branch_Census**
Rank (Officer/Enlisted)	P	**Officer_Census**
State	A	

A = actual data   **fully adjusted**
P = imputes with   (partially adjusted)
     parameter

After the April 1, 2000, baseline has been generated, the following steps are necessary to create subsequent years' data (for September 30, 2000, through September 30, 2006 – the latest year for which data were available at the time the model was developed). We determine the number of veterans in the next year (time t, age a) as shown in the equation below:

- start with the number of veterans in the previous year (for example, 24 year olds in September 2000)

- add new separations

- subtract deaths

- account for interstate migration

$$\text{Vets}_{t,a} = \text{Vets}_{t-1,\ a-1} + \text{seps}_{fy(t)} - \text{Deaths}_{fy(t)} +/- \text{Migration}_{fy(t)}$$

t-1, a-1		t,a
9/00	**fy01**	**9/01**
24 yrs	*Separations, Deaths*	*25 yrs*
Living	*and Migration*	*Living*

DMDC data are used to count veterans who separated between April 1, 2000, and September 30, 2006. We begin processing the file by removing "out-of-scope" records (that is, records of those who are currently active or who separated from active duty prior to April 1, 2000, or after September 30, 2006). We then use parameters to fill in characteristics such as Race and State of residence (which are not available from DMDC).

As the last step, we combine the new separations with the existing veterans before applying mortality and migration. The actual death data from DMDC is incorporated, while a parameter is used to determine the mortality status of those veterans who are included in the Census data. We use the same migration parameter for all veterans to account for state-to-state migration.

**GENERATING
SEPTEMBER 30, 2000 – SEPTEMBER 30, 2006**

	Census		DMDC	
Separations	X		A	
Deaths	P	Mort_Census	A	
Migration	P	Migrate_Historical	P	Migrate_Historical

X = does not apply
A = actual data
P = imputes with parameter

**NEW SEPARATIONS BETWEEN
APRIL 2, 2000 – SEPTEMBER 30, 2006**

	DMDC	
Category	**Parameter**	**Data Adjustment**
Age	A	
Gender	A	
Period of Service	A	
Living	A	
Race	A	
Veterans	A	
Branch	A	(Vets_Dmdc)
Rank (Officer/Enlisted)	A	
State	P	**State_Seps**

A = actual data   **fully adjusted**
P = imputes with parameter   (partially adjusted)

Projection

The projection phase creates data for September 2007 and forward, as described in the equation above, but often with different parameters.

Projections of future separations are based on GORGO data provided by DoD's OACT. Some variables, such as Gender, have to be filled in with parameters; these apply to assumptions needed to transition into the future, such as deaths and migration.

PROJECTING BEYOND SEPTEMBER 30, 2006

	Census-DMDC		GoRGo
Separations	X		E
Deaths	P	Mort	E
	P	Mort_Adjust	
Migration	P	Migrate_Future	P

X = does not apply
E = GORGO estimate
P = imputes with parameter

PROJECTED SEPARATIONS AFTER SEPTEMBER 30, 2006

	GoRGo	
Category	**Parameter**	**Data Adjustment**
Date	E	
Age	E	
Gender	P	**Gender_DoD**
Period of Service	E	
Living	E	
Race	P	**Race_DoD**
Veterans	E	(Vets_DoD)
Branch	P	**Branch_DoD**
Rank (Officer/Enlisted)	E	
State	P	**State_DoD**

E = GORGO estimate
P = imputes with parameter

fully adjusted
(partially adjusted)

For example . . .

Here are some examples to help understand how the programs work:

- **A veteran from the Vietnam era** would be accounted for through Census data, which would provide aggregate information on Age, Gender, Period of Service, Race, and State. Parameters fill in information on Branch and Rank. This information would then form a basis for the April 2000 estimate. Each veteran would be aged, and from that point on, parameters would supply annual probabilities for migration and mortality.

- **An Iraqi war veteran** who separated in early 2002 would be represented in the 2002 estimate by DMDC data, providing person level information on Age, Gender, Rank, and Branch. Parameters fill in information on State and Race.

  In subsequent estimates, this veteran is aged and mortality status is included, using actual data.

  A parameter is applied to represent the annual migration, as this information is not available after separation. In projections after 2006, actuarial assumptions are applied via parameters to represent annual probabilities for mortality and migration.

- **A post-2006 veteran,** on the other hand, would be represented in the GORGO data, which would indicate information on Age, Rank, and Period of Service. Parameters are used to estimate Gender, Race, Branch, and State. Parameters based on annual probabilities would determine mortality and migration.

In summary, the estimate for 2007 would include some veterans from the Vietnam era through the 2000 Census data; some veterans (April 2000 through September 2006 separations) from DMDC data; and some newly separated veterans (after 2006) from GORGO data.

Output Files

The VetPop2007 model stores important model output in Microsoft Access databases, from which a Visual Basic for Access (VBA) program creates a series of Microsoft Excel Pivot Tables. These tables give users easy access to tabulations of model output, offer flexibility in the choice of row and column variables, and allow the restriction of tabulations to certain subgroups of veterans. A list of Excel reports is provided in Overview Table 2. below.

Non-technical users can easily manipulate the Excel reports, which are based on a subset of the information in the SAS data sets. Accessing the SAS data sets requires the use of SAS Version 8 or higher.

Note: some reports (e.g., 4D, 4s, 4M, 5D, and 5s) are not produced because of limitations in the model, but they are options for a future VetPop model.

OVERVIEW I TABLE 2. VETPOP2007 STATE AND NATIONAL TABLES

Row	Column	Pages	Living	Deaths	Separations
State	Age (5-yr)	Year, Gender, Period of Service	1L	1D	1S
State	Period of Service*	Year, Gender, Age (5-yr)	2L	2D	2S
State	Race/Ethnicity	Year, Gender, Age (5-yr)	4L		
Year	Race/Ethnicity	Age (1 -yr), Gender, Period of Service	5L		
Year	Branch of Service	Age (1 -yr), Gender, Period of Service	8L	8D	8S
Year	Officer/Enlisted	Age (1 -yr), Gender, Period of Service	9L	9D	9S
Year	Gulf War Division	Age (1 -yr), Gender, Period of Service	10L	10D	10S

* Differs from the "Period of service" variable appearing in all other reports, in that it contains subtotals of certain periods.

System/Software

The great majority of the work was done in SAS in a PC environment, except for a few parameters developed in Stata. The final process is to convert the SAS data sets into small Microsoft Access files, which are imported into Excel Pivot tables, using a VBA Program. The pivot tables are then posted to the VA Internet (go to http://www1 .va.gov/vetdata/ and click on Demographics).

You can request CDs via e-mail at VETPOP@VA.GOV. OACT can also provide SAS and Access files to appropriate researchers.