Appendix A.
Statistical Methodology

THE CENSUS POPULATION

The 2014 Census of Horticultural Specialties (CHS) was designed to cover all operations from which $10,000 or more of horticultural products were produced and sold, or normally would have been sold, during 2014. Horticultural products include bedding plants, potted flowering plants, cut flowers, cut cultivated florist greens, trees, shrubs, ground covers, vines, fruit and nut trees, sod, dry bulbs, greenhouse produced vegetables, commercial vegetable transplants, vegetable and flower seeds, Christmas trees, short term woody crops, aquatic plants, unfinished or prefinished plants, propagation materials, and other nursery or greenhouse plants.

Data collection for the 2015 Commercial Floriculture Survey was conducted in conjunction with the 2014 Census of Horticultural Specialties. Supplemental questions, not summarized in the 2014 Census of Horticultural Specialties, were included in the data collection to meet the requirements needed for the 2014 Floriculture Crops Summary report.

The 2014 CHS mail list was built from NASS's list frame. All records on the frame with $10,000 or more in horticultural sales were included on the mail list. A sample was selected for other horticultural operations on the frame that had less than $10,000 in horticultural sales or had unknown sales values. The final mail list included 40,319 operations. The response rate is an indicator of quality of data collection. The response rate for the 2014 CHS was 60 percent and is calculated by the ratio of completed records with $10,000 or more of horticultural sales divided by the sample excluding records with less than $10,000 in horticultural sales.

DATA COLLECTION

Method of Enumeration

The 2014 CHS was accomplished primarily by mailout/mailback, but supplemented with Electronic Data Reporting (EDR) on the Internet, telephone enumeration, and personal enumeration for special classes of records. Personal enumeration (interviewing) involved the use of both Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI). Office enumerators at the NASS National Operations Division (NOD) in St. Louis, Missouri, with assistance from NASS staff in Montana and Arkansas, conducted CATI data collection. In addition field enumerators conducted phone and personal interviews with respondents. For the 2014 CHS, NASS implemented a pre-notification strategy in an effort to increase awareness, improve overall responses, and encourage respondents to report early to avoid continued correspondence. All records in the initial mailout received either a postcard or pre-recorded voice message announcing the census mail packets were coming.

Report Form

Three 28-page report forms were used to capture the number of horticultural products produced and sold and the value of sales for both retail and wholesale sales. The various types of plants sold were grouped by sections in the report forms.

The three report forms used for the 2014 CHS included a U.S. (excluding Hawaii) horticulture report form (14-A0624), a U.S. (excluding Hawaii) floriculture report form (14-A0625), and a Hawaii horticulture report form (14-A0627). The U.S.

horticulture report form and the U.S. floriculture report form were exactly the same with the exception that they were printed in different colors to differentiate between horticulture operations (green forms) and floriculture operations (yellow forms). The Hawaii horticulture report form content was unique. All of the report forms allowed respondents to write in specific commodities that were not listed on their form.

Additionally, information was obtained for area in production for several types of crops; marketing channels; estimated value of land, buildings, machinery, and equipment; production expenses; and the number of hired workers employed by the operation in 2014. See Appendix B for facsimiles of the report form and instruction sheet.

Report Form Mailings

NASS's North Carolina Print Mail Center (PMC) began pre-notification by postcard on December 1, 2014. The 2014 CHS report form was mailed from the Census Bureau's National Processing Center (NPC) at Jeffersonville, Indiana on December 15, 2014. Each operation selected for the census was mailed a packet that contained a cover letter, an EDR instruction letter, a report form instruction sheet, a labeled report form, and a return envelope addressed to either NPC or NOD for data capture. The report form carried a return due date of February 5, 2015. NPC was contracted to perform mail packet preparation, initial mailout, and follow-up mailing to nonrespondents.

The follow-up mailing took place from NPC on February 19, 2015.

Personal Follow-up

Telephone follow-up interviews to non-respondents took place from March 2 to July 10, 2015 from a NASS Data Collection Center.

Data collection for the 2014 CHS was coordinated with other NASS surveys. In some cases, if a horticultural operation was also selected for a survey, NPC mailed the 2014 CHS materials to NASS Regional/Field Offices. Office personnel were responsible for collecting the horticulture data and completing other survey report forms in the most efficient way to reduce the number of contacts and minimize respondent burden.

REPORT FORM PROCESSING

Data Capture

NPC received and processed returned mail packets for the U.S. horticulture and floriculture report forms (14-A0624 and 14-A0625). NASS staff on site at NPC provided technical guidance and monitored NPC processing activities. All report forms returned to NPC were immediately checked in, using bar codes printed on the mailing label, and removed from follow-up report form mailings. All forms with any data were scanned and an image was made of each page of a report form. Optical Mark Recognition (OMR) was used to capture categorical responses and to identify the other answer zones in which some type of mark was present.

All forms were reviewed prior to data keying to identify inconsistencies and ensure the data could be keyed. Major inconsistencies, respondent remarks, and blank forms were reviewed by analysts and adjusted prior to keying. In some cases, report forms were mailed to regional field offices for further editing. All forms with any data were scanned and an image was created for each page of the report form.

Data entry operators keyed data from the scanned images using OMR results that highlighted the areas of the report forms with respondent entries. The keyer evaluated the contents and captured pertinent responses. Ten percent of the captured data were keyed a second time for quality control. If differences existed between the first keyed value and the second, an adjudicator handled resolution. The decision of the adjudicator was used to grade the performance of the keyers, who were required to maintain a certain accuracy level.

The images and the captured data were transferred to NASS's centralized network and became available to regional field offices and headquarters on a flow basis. The images were available for use in all stages of review. Images were computer generated for reports obtained from the telephone interviews and the Internet.

The NOD processed returned mail packets for all of

the Hawaii horticulture report forms (14-A0627). All forms with any data were keyed from image, scanned, and an image was created for each page of the report form.

Editing Data

Captured data were processed through a computer formatting program, which verified that records were valid. Rejected records were referred to analysts for correction. Accepted records were sent to a complex computer batch edit process. Each execution of the computer edit in batch mode consisted of records from only one State and flowed as the data were received from each data collection source.

The computer edit determined whether a reporting operation met the qualifying criteria to be counted as an in-scope record. The edit examined each in-scope record for reasonableness and completeness and determined whether to accept the recorded value for each data item or to take corrective action. Such corrective actions included removing erroneously reported values, replacing an unreasonable value with a value consistent with other reported data, or providing a value for an overlooked item. To the extent possible, the computer edit determined a replacement value. Strategies for determining replacement values are discussed in the next section. Operations failing to meet the qualifying criteria were categorized as out-of-scope. Out-of-scope records that NASS had reason to believe might be in-scope (indications of recent and/or significant horticultural activity reported on NASS surveys, for example) were referred to analysts for verification.

The edit systematically checked reported data section-by-section with the overall objective of achieving an internally consistent and complete report. NASS subject-matter experts had previously defined the criteria for acceptable data. Problems that could not be resolved within the edit were referred to an analyst for intervention. Regional and field office analysts also participated using an interactive version of the edit program to submit corrected data and immediately re-edit the record to ensure satisfactory resolution.

In some cases, respondents may have failed to provide all of the information requested, only indicating the presence of an item but not the amount. These items were coded for computer imputation.

After the initial edit, an automated imputation program supplied missing data based on state or national averages. A post-imputation computer edit was performed to ensure imputation actions provided acceptable results. Instances where imputed data failed edit checks were referred to analysts for corrective action.

Data Analysis

The complex edit ensured the full internal consistency of the record. Successfully completing the edit did not provide insight as to whether the report was reasonable compared to other reports in the county. Analysts were provided an additional set of tools to review record-level data across operations. These examinations revealed extreme outliers, large and small, or unique data distribution patterns that were possibly a result of reporting, recording, or handling errors. Potential problems were researched and, when necessary, corrections were made and the record interactively edited again.

ESTIMATION

Nonresponse Weighting

The 2014 Census of Horticultural Specialties is a census of every operation on the NASS Horticulture Sampling Frame with at least $10,000 of horticultural sales indicated. Operations on the frame that had indicators of horticultural sales below the $10,000 threshold were sampled at an average rate of 1 out of 8.

Although much effort was expended to obtain a response from each operation selected for the census, it was not possible to obtain a complete set of responses. Nonresponse can lead to biases in published estimates because the information concerning the horticultural enterprise production on the nonresponding operations could not be factored into the estimates. Such estimates of totals will be biased low. To reduce this bias, NASS made nonresponse adjustments to the initial weights of the responding operations. The nonresponse weight adjustment increases the weight of responding operations to account for the data that would have

been reported by the nonresponding operations. This increased the estimates of totals obtained by the respondents and reduced this bias.

Conceptually, each operation on the sample begins the weighting process with an initial weight equal to the inverse of the record's probability of selection. Records with sales of $10,000 or more will have an initial weight of 1 because they are selected with certainty. Records with sales less than $10,000 will have an initial weight of about 8.
If each operation selected for the census provided the requested data, the data could simply be multiplied by each record's initial weight then added up to attain an estimate for the total amount of the item of interest. In the presence of nonresponse, nonresponse adjustments are computed and applied to the initial weights of the responding operations resulting in a nonresponse-adjusted weight greater than the initial weight for these operations. The initial weight of each nonresponding operation is then adjusted to zero. The adjustments are computed in a manner that requires the sum of the nonresponse-adjusted weights across the responding operations on the survey to equal the number of records on the sampling frame.

Nonresponse Weight-Adjustment Groups

To compute nonresponse adjustments, each operation on the mail list was placed in a weight-adjustment group. Each operation was assigned to a group based on the characteristics used to define the group. It was necessary that the characteristics that defined the weight-adjustment groups were available for responding and nonresponding operations alike. Therefore, it was not possible to define weight-adjustment groups using data collected via the CHS.

The information on the sampling frame was used to create the weight-adjustment groups and was a measure of the horticultural economic size (HES). The basic definition of the weight-adjustment groups is given below:

Definition

HES<$10,000
$10,000<=HES<$50,000
$50,000<=HES<$150,000
$150,000<=HES<$250,000
$250,000<=HES<$500,000
$500,000<=HES
Must Group (varies by state)

All records that were considered likely to be very large horticultural operations for a given state where considered "must" cases and put in a special group. For all records in a must group, nonresponse adjustment was not allowed and data were imputed for any of these records that did not respond. Must group definitions varied by state.

Nonresponse-Adjustment Computation

A separate nonresponse adjustment was calculated within each weight-adjustment group. All responding records within each group received the same nonresponse adjusted weight. The nonresponse-adjustment was obtained by dividing the sum of the initial weights across all the records in the group by the sum of the initial weights of the responding operations in the group. If the sum of the initial weights across all records in the group was 50 and the sum of the initial weights of all responding operations in the group was 40, the nonresponse-adjustment for the responding operations was 50/40 or 1.25. The nonresponse-adjusted weight for all responding operations in the group was the product of the initial weight and the nonresponse adjustment of 1.25. This was simply (1 x 1.25). Note that 1.25*40=50, the sum of the initial weights for all records in the group.

The assumption made when computing nonresponse adjustments in this way was that within each weight-adjustment group, the data that the nonrespondents would have provided had they responded were collectively similar to the data provided by the respondents. This assumption was made somewhat more plausible because operations in the same group shared similar characteristics with respect to the information used to define the group - the HES.

Accounting for Misclassification

When conducting censuses, it is possible that respondents might inadvertently report some data in error. Operations that really should be determined to be in scope for the CHS, i.e., have at least $10,000 worth of horticultural sales, might report on the CHS that their horticultural sales are less than the

threshold. Conversely, operations that report that they meet the threshold on the CHS might in fact not actually meet it and should be considered out of scope for the census.

In order to measure the impact of misreporting scope status, NASS conducted a misclassification survey that consisted of a small sample of CHS respondents. A small set of screener questions was asked to determine the true scope for each of the operations selected for the misclassification survey. Using this methodology, misclassification adjustments were computed and used to adjust the nonresponse weights of the CHS respondents to account for reporting errors with respect to CHS scope status.

Coverage Weighting Adjustments

The target population for the 2014 CHS was all operations that had at least $10,000 of commercial horticultural production in 2014. Unfortunately, it is impossible to compose a list of operations that is complete. Due to this incompleteness of the mail list, data produced from it, even if perfectly corrected to account for nonresponse, will still have a tendency to be biased downwards because operations not on the list would not have any representation. This bias due to list incompleteness is called coverage bias, or more specifically, bias due to undercoverage of the sampling frame.

To reduce the amount of this bias, an additional adjustment was calculated and applied to the nonresponse-adjusted weight for each responding operation. This was called the coverage adjustment.

Coverage Adjustment Computation

The majority of CHS respondents were also respondents on the 2012 Census of Agriculture. Operations that were respondents to both censuses were assigned the census of agriculture coverage adjustment computed for the operation in the 2012 Census of Agriculture. The coverage adjustment for CHS respondents that did not match the census of agriculture were calculated using records with similar information that did match the census of agriculture.

The coverage adjustment was then applied to the misclassification-adjusted nonresponse weight for each CHS respondent record. This resulted in a fully-adjusted weight. The fully-adjusted weight attempts to correct for nonresponse and misclassification bias, as well as coverage bias.

Summary Weights

Most of the fully-adjusted weights for the 2014 Census of Horticultural Specialties were not whole numbers (integers). Using these weights to create the estimates published in the tables would result in fractional values. These would be difficult to read and cause consistency problems between related tables. To avoid some of these problems, summary weights were created by randomly moving the fully-adjusted weights up or down to an integer in a way that preserved the overall sum of the fully adjusted weights. This process is called weight integerization. The resulting summary weights were used to produce the numbers published in the tables.

MEASURES OF PRECISION AND ACCURACY OF THE ESTIMATES

All numbers published in the tables are estimates of particular characteristics of the entire population of horticultural operations. The true values of these characteristics are unknown and unknowable. Even though an attempt was made to obtain a response from every operation selected for the survey and weight adjustments computed, the data produced by the census will not attain the true values. This is due to the fact that weight adjustments are imperfect and the assumptions on which those adjustments are made are imperfect as well. Hypothetically, if the entire census process was repeated over and over again, each replication of the census would almost certainly produce a different result for the same true population value every time. This is because each time the census is carried out, a different set of respondents would be obtained, response rates would fluctuate, and calculated weight adjustments would not be exactly the same.

It is possible to obtain an idea of how much this variation would be on average by calculating the estimate's variance. The estimated variance of an estimate gives a measure of the average squared random fluctuation that would be seen in an estimate if the census was carried out multiple times. Because the variance measures random fluctuation in squared

units, the square root of the variance is computed to obtain a random fluctuation measure that is in the same units as the original estimate. This is called the standard error (se) of the estimate. The standard error can then be divided by the estimate itself to show the relative size of the standard error to the estimate. This ratio is known as the coefficient of variation. If this ratio is small, the estimate is quite precise. If this ratio is large, the estimate is imprecise. An estimate of 100 with a standard error of 2 would result in a relative standard error of .02 or 2 percent. This would be a very precise estimate. An estimate of 100 with a standard error of 30 would result in a relative standard error of 30 percent. This might be considered to be an imprecise estimate. The idea of precision can be made a little more clear by stating that if the estimate is 100 with a standard error of 2, you could be quite confident that the true population value would be in the interval 96 to 104 (within two standard errors of the estimate).

Table A provides statistical precision estimates for the number of farms, total sales, wholesale sales, retail sales for the United States and for each state. Table B provides statistical precision estimates for the total value of sales by size and operations by type of crop for the United States.

Table A. **Reliability Estimates of Operations and Value of Sales of All Horticultural Specialty Crops – United States and States: 2014**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	Operations		Total sales		Wholesale sales		Retail sales	
	Number	Coefficient of variation (percent)	Value ($1,000)	Coefficient of variation (percent)	Value ($1,000)	Coefficient of variation (percent)	Value ($1,000)	Coefficient of variation (percent)
United States	23,221	1.1	13,789,048	0.9	11,859,976	1.0	1,929,073	2.3
Alabama	280	5.9	249,304	1.9	235,384	2.6	13,921	18.6
Alaska	65	16.1	16,832	78.4	3,858	97.5	12,974	81.4
Arizona	155	9.0	388,893	11.2	328,062	13.7	60,831	6.3
Arkansas	104	10.2	34,390	7.8	24,741	17.1	9,649	28.4
California	1,710	4.8	2,877,981	2.3	2,625,696	2.9	252,286	7.3
Colorado	253	6.6	201,897	6.2	150,549	7.0	51,349	9.6
Connecticut	492	27.2	286,732	1.8	244,625	0.7	42,108	9.7
Delaware	51	17.1	21,774	52.4	12,819	90.3	8,955	5.3
Florida	2,069	4.1	1,796,747	1.3	1,699,047	1.8	97,700	14.2
Georgia	431	5.9	305,575	4.6	271,695	5.3	33,880	10.0
Hawaii	450	1.4	81,204	10.1	68,377	9.7	12,826	13.2
Idaho	273	9.0	67,418	10.4	49,384	9.7	18,034	34.7
Illinois	519	4.9	258,320	5.1	196,292	8.6	62,028	20.7
Indiana	434	4.9	135,484	4.3	104,754	5.0	30,730	12.4
Iowa	289	4.7	99,330	6.5	73,819	6.3	25,512	10.8
Kansas	169	13.1	46,981	7.3	29,193	12.2	17,788	22.6
Kentucky	473	9.6	83,097	6.7	60,374	9.1	22,723	20.4
Louisiana	259	20.2	84,546	14.0	66,249	5.6	18,297	51.8
Maine	320	5.0	70,544	10.6	22,665	29.6	47,878	5.9
Maryland	313	5.4	225,968	2.3	207,368	3.1	18,600	11.8
Massachusetts	583	5.4	102,946	5.1	59,499	9.2	43,447	9.9
Michigan	940	1.8	644,981	1.1	545,425	1.1	99,556	3.2
Minnesota	487	5.9	244,618	6.8	168,755	6.5	75,862	11.6
Mississippi	180	15.1	44,311	13.9	30,340	18.2	13,971	16.7
Missouri	291	7.4	84,826	9.4	64,187	8.9	20,639	24.2
Montana	128	11.5	24,321	18.7	11,936	18.8	12,385	30.8
Nebraska	148	8.3	48,470	11.1	25,912	25.2	22,558	17.9
Nevada	18	38.4	33,286	84.4	30,212	95.2	3,074	46.1
New Hampshire	212	12.8	43,275	4.5	28,446	3.8	14,829	8.2
New Jersey	630	5.3	355,730	5.1	314,123	6.0	41,607	7.3
New Mexico	92	19.8	32,328	9.7	22,857	7.2	9,471	23.8
New York	1,058	2.8	324,129	1.7	250,456	1.7	73,673	2.0
North Carolina	1,337	5.0	570,686	2.7	506,695	2.8	63,991	7.6
North Dakota	37	27.0	6,560	27.8	3,391	33.0	3,169	38.7
Ohio	778	2.8	392,065	1.6	313,029	1.6	79,036	2.9
Oklahoma	146	7.7	149,781	7.4	126,989	8.8	22,792	18.0
Oregon	1,281	2.8	932,041	2.2	875,225	2.2	56,816	8.6
Pennsylvania	1,397	2.7	326,641	4.7	238,768	5.9	87,873	4.5
Rhode Island	155	11.7	20,354	43.5	15,006	55.2	5,348	28.2
South Carolina	250	10.5	151,038	3.9	132,512	2.7	18,527	15.7
South Dakota	62	11.6	17,911	10.9	9,909	16.0	8,002	21.2
Tennessee	547	3.6	256,507	2.9	204,158	3.0	52,349	8.1
Texas	682	3.4	593,779	3.7	533,768	4.0	60,011	8.1
Utah	125	10.6	86,044	4.3	67,571	5.0	18,473	28.3
Vermont	281	18.1	24,738	10.8	7,016	14.9	17,722	13.7
Virginia	560	5.5	323,255	3.7	296,933	4.0	26,322	11.4
Washington	709	3.2	365,679	16.9	326,536	13.6	39,142	44.3
West Virginia	131	12.9	21,021	3.1	14,209	3.0	6,812	9.2
Wisconsin	833	2.9	230,693	8.5	160,261	10.8	70,431	10.5
Wyoming	34	29.5	4,018	55.8	901	(H)	3,116	61.2

Table B. **Reliability Estimates of Operations and Value of Sales for Selected Horticultural Specialty Items – United States: 2014**

[For meaning of abbreviations and symbols, see introductory text.]

Item		Total	Coefficient of variation (percent)
TOTAL VALUE OF SALES BY SIZE			
$10,000 - $19,999	operations	4,203	6.7
	$1,000	59,489	7.4
$20,000 - $24,999	operations	1,458	5.9
	$1,000	32,380	6.0
$25,000 - $39,999	operations	2,796	3.7
	$1,000	88,385	3.5
$40,000 - $49,999	operations	1,334	3.7
	$1,000	59,490	3.6
$50,000 - $99,999	operations	3,709	3.0
	$1,000	261,655	3.1
$100,000 - $249,999	operations	3,729	1.5
	$1,000	586,303	1.6
$250,000 - $499,999	operations	2,098	3.4
	$1,000	733,822	4.0
$500,000 - $999,999	operations	1,582	3.5
	$1,000	1,114,059	3.8
$1,000,000 - $2,499,999	operations	1,344	2.5
	$1,000	2,082,590	2.8
$2,500,000 or more	operations	968	1.6
	$1,000	8,770,877	1.6
OPERATIONS BY TYPE OF CROP			
Annual bedding/garden plants	operations	7,964	1.3
	$1,000	2,567,534	2.0
Potted herbaceous perennial plants	operations	6,291	2.5
	$1,000	944,850	0.8
Potted flowering plants for indoor or patio use	operations	4,059	2.2
	$1,000	1,084,274	1.3
Foliage plants for indoor or patio use	operations	2,644	4.8
	$1,000	721,889	2.3
Cut flowers and cut lei flowers	operations	1,998	2.6
	$1,000	462,098	5.1
Cut cultivated greens	operations	728	8.3
	$1,000	99,040	1.5
Nursery stock sold (see text)	operations	8,226	2.1
	$1,000	4,266,631	1.2
Propagative horticultural materials, bareroot, and unfinished plants (see text)	operations	1,067	5.5
	$1,000	695,126	1.3
Sod, sprigs or plugs sold (see text)	operations	1,289	4.2
	$1,000	1,138,465	2.8
Dried bulbs, corms, rhizomes, and tubers	operations	204	7.9
	$1,000	74,014	26.3
Food crops grown under protection	operations	2,521	4.0
	$1,000	796,664	8.0
Transplants for commercial vegetable and strawberry production	operations	693	11.4
	$1,000	371,817	2.5
Vegetable seeds	operations	385	10.0
	$1,000	135,122	55.9
Flower seeds	operations	169	21.1
	$1,000	31,607	9.9
Aquatic plants	operations	345	11.2
	$1,000	20,205	13.6
Cut Christmas trees sold (see text)	operations	3,352	6.7
	$1,000	366,632	2.8
Short rotation woody crops sold (see text)	operations	49	29.4
	$1,000	1,845	30.7
Tobacco transplants	operations	183	9.5
	$1,000	11,236	19.3