An Evaluation of Persons Per Household (PPH) Data Generated By the American Community Survey: The Demographic Perspective

David A. Swanson
University of California Riverside
George C. Hough, Jr.
Portland State University

ACS Multi-Year Studies Meeting
US Census Bureau
Suitland, MD
November 15, 2007

OUTLINE

- INTRODUCTION
- DATA
- RESULTS
- DISCUSSION
- CONCLUSIONS & FUTURE RESEARCH

INTRODUCTION

Does the ACS produce PPH (Persons Per Household) estimates that applied demographers will use?

1. The American Community Survey (ACS) provides annual estimates of persons per household (PPH) that are subject to sample error. This means that they can fluctuate from year to year in a given population, which reflects a "Statistical Perspective."

2. Demographers tend to view PPH as a population attribute that has demographic determinants.

3. This implies that demographers view PPH as an attribute that changes systematically over time (the "Demographic Perspective"), not randomly (the "Statistical Perspective").

INTRODUCTION

Does the ACS produce PPH (Persons Per Household) estimates that applied demographers will use? (Cont.)

1. Applied demographers need PPH in making population estimates with the housing unit method (HUM).

2. Do PPH values generated by the ACS values fit sufficiently into the demographic perspective that they are likely to be used by applied local demographers in conjunction with the HUM?

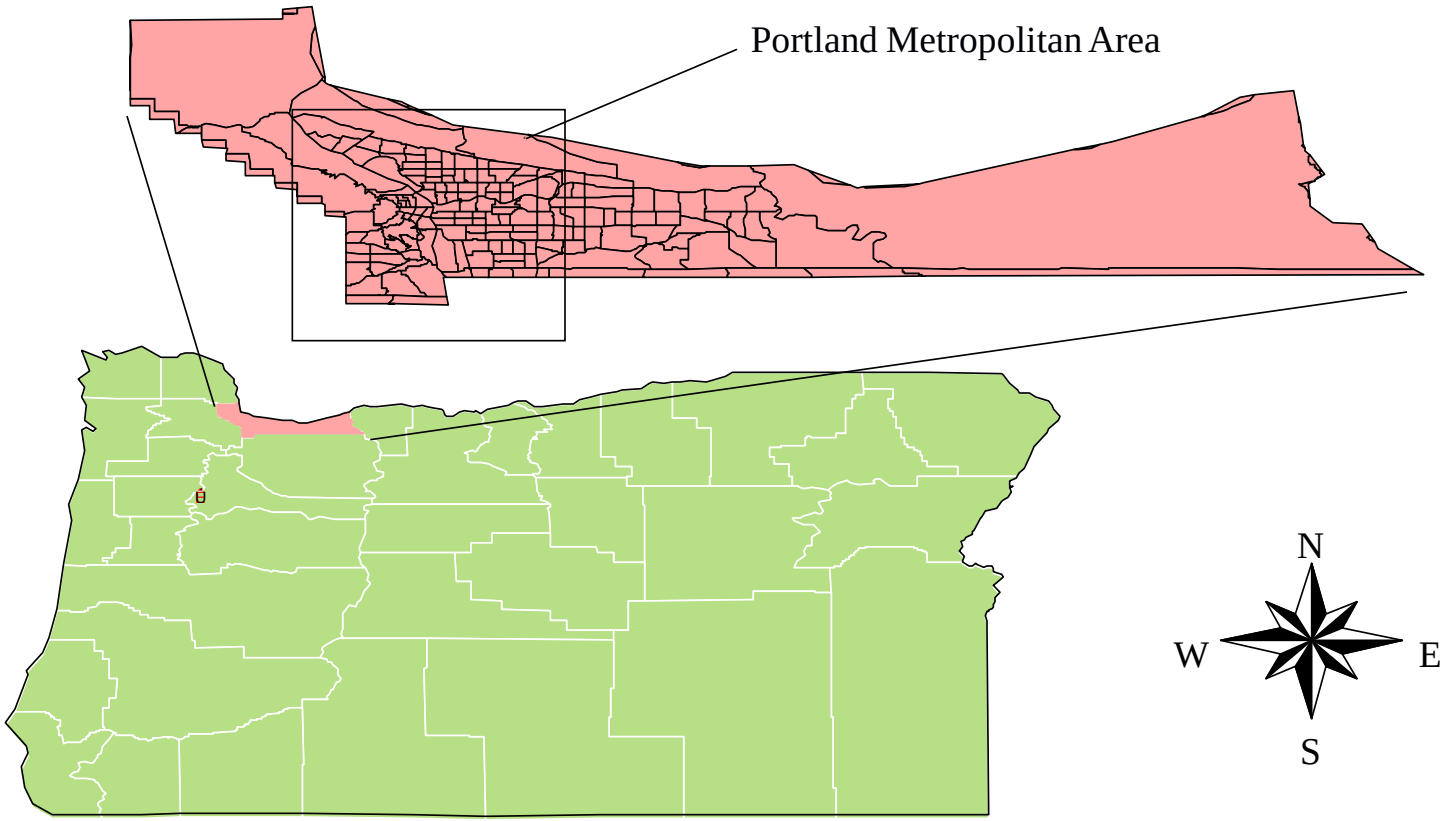
INTRODUCTION

UCR University of California, Riverside

What generated this research?

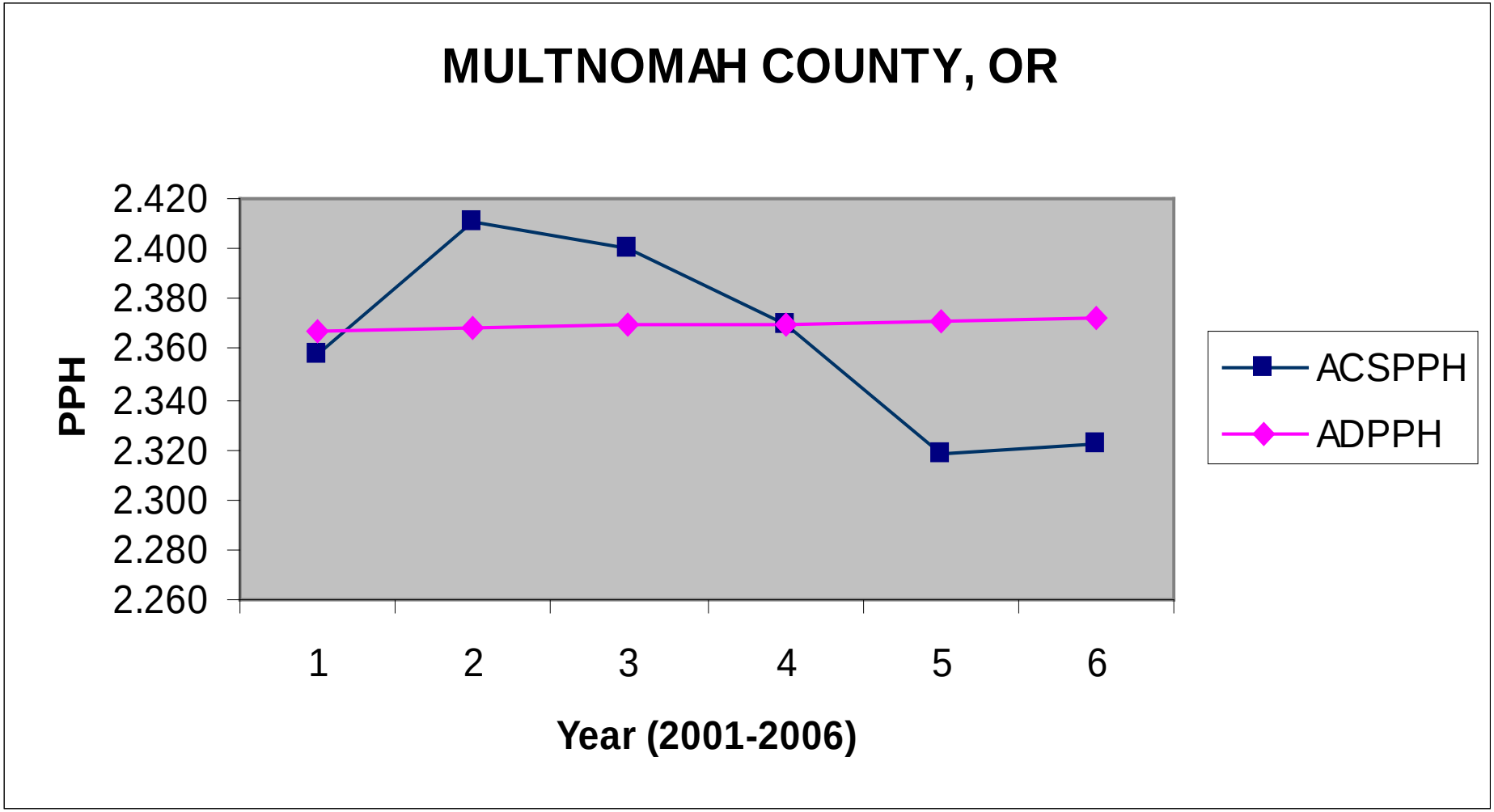
We were looking at the Multnomah County PPH values from the ACS for 2001 to 2006 and saw some “jumping around,” which got our attention as demographers who use the housing unit method to do estimates. We then compared this to a set of PPH values generated by extrapolating the change in PPH between the 1990 and 2000 censuses for Multnomah County.

INTRODUCTION



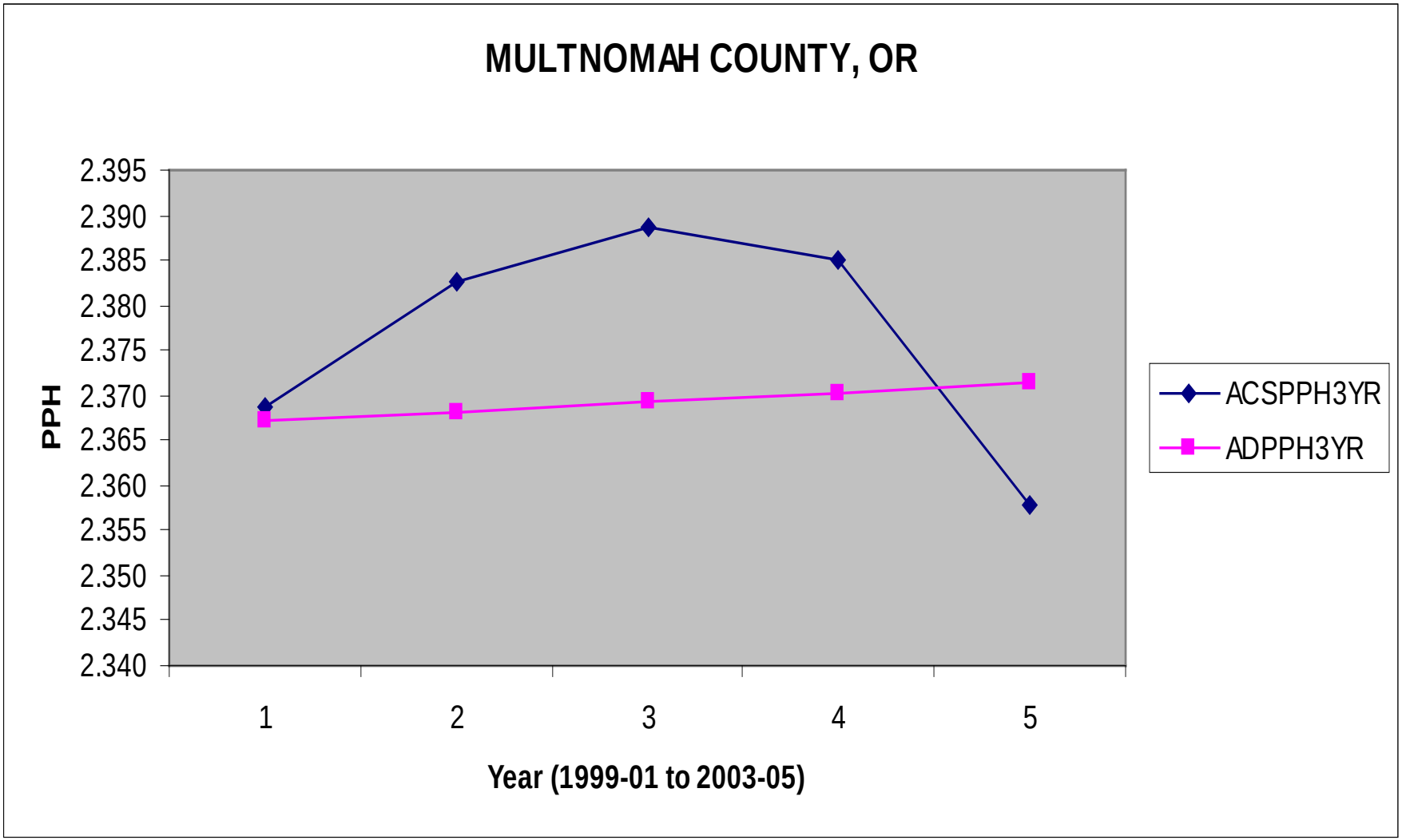
INTRODUCTION

ACS PPH v. AD PPH: SINGLE YEAR ACS DATA



INTRODUCTION

ACS PPH v. AD PPH: THREE YEAR ACS DATA



INTRODUCTION

The formula used in the Housing Unit Method (HUM) to generate a population for a given point in time is

$$P = GQ + (PPH)(H)(OR)$$

where

P = Total Population

GQ = Population in Groups Quarters

PPH = Persons Per Household

H = Total Number of Housing Units

OR = Occupancy Rate

Note that (H)(OR) = Total Number of Households

INTRODUCTION

These differences are of interest because they lead to very different estimates of household populations. For example, using the same housing unit counts and occupancy rates to get the number of occupied units (303,086) estimated by the ACS for 2005, the difference between using a PPH of 2.319 (the ACS estimate for Multnomah County in 2005) vs. a PPH of 2.371 (the analytically derived estimate for Multnomah County in 2005), we find an estimated household population of 656,146 for Multnomah County in 2005 using the ACS PPH value and an estimated household population of 671,056 using the PPH value derived analytically.

INTRODUCTION

UCR UNIVERSITY OF CALIFORNIA, RIVERSIDE

This is a tremendous difference for states that use population estimates for revenue sharing. For example, at $200 per person, the ACS estimate provides Multnomah County with $2,982,000 more than does the alternative (14,910 * $200). Because state revenue sharing is a zero sum process, this amount would come from what otherwise would have been distributed to Oregon's 35 other counties. This led us to examine the other 20 ACS test sites and this examination forms the basis of our presentation today.

DATA

The 18 (of 36) 1999 ACS test counties used in this study

Pima County, AZ
Jefferson County, AR
San Francisco County, CA
Tulare County, CA
Broward County, FL
Lake County, IL
Black Hawk County, IA
Calvert County, MD
Hampden County, MA
Madison County, MS
Douglas County, NE
Bronx County, NY
Rockland County, NY
Franklin County, OH
Multnomah County, OR
Schuylkill County, PA
Sevier County, TN
Yakima County, WA

DATA

For these counties we assembled PPH values from the American Community Survey as follows:

(1) Single Year Estimates, 2001 to 2006; and
(2) Three-Year Estimates, 1999-2001 to 2003-2005

To compare against PPH values derived analytically.

DATA

The analytical method for generating PPH in this study is a method commonly used by applied demographers for this purpose, namely, the geometric rate of change benchmarked to two most recent successive census counts and applied to the PPH value found in the most recent census count, which is then extrapolated beyond the most recent census by applying the rate of change to it.

DATA

The process takes place in two steps. The first is the calculation of the rate of change in PPH:

$$r = (PPH_l/PPH_b)^{(1/y)} - 1$$

where

r = rate of change

PPH = Persons Per Household

l = Launch Year (most recent census)

b = Base Year (Census preceding launch Year

y = Number of years between l and b (10 years)

DATA

The second step is applying the rate to the launch year to find PPH values:

$$PPH_t = (PPH_l)\,[(1 + r)^{(y)}]$$

where

r = rate of change (from step 1)

PPH = Persons Per Household

t = Target Year

l = Launch Year (most recent census)

y = number of years between t and l

DATA

To demonstrate that this method works well, here is the result of a test in which estimated PPH values in 2000 are compared to Census 2000 PPH values for the 39 counties of Washington.

The estimated PPH values in 2000 are based on applying the changes in PPH values observed between the 1980 census and 1990 census to 1990 census PPH values.

DATA

UCR University of California, Riverside

Accuracy Test of the Geometric Method of Estimating PPH Values

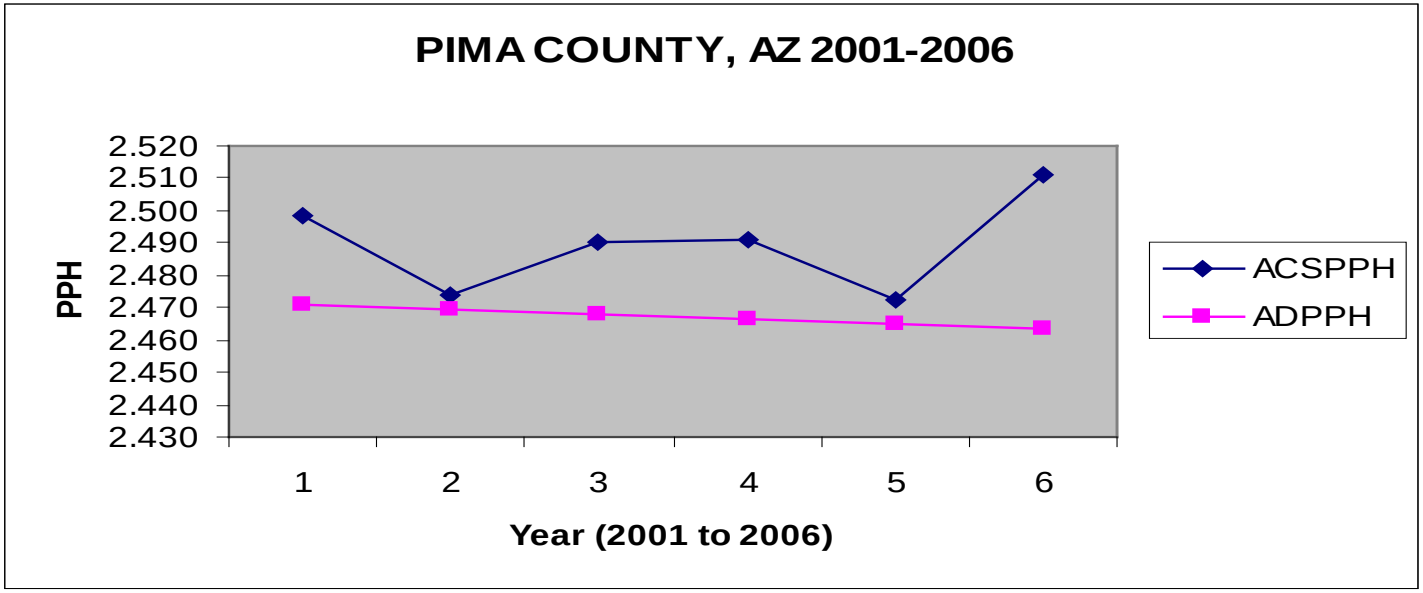
	1980 Persons Per Household	1990 Persons Per Household	2000 Persons per Household	1980-1990 Geometric Rate of Change	Estimated 2000 Persons Per Household	Absolute Error	Percent Error	MAPE
STATE	2.6086	2.5348	2.5349	-0.0029	2.4631	-0.0718	-2.83%	2.83%
Adams	2.9113	2.9405	3.0949	0.0010	2.9700	-0.1249	-4.03%	4.03%
Asotin	2.5662	2.4727	2.4162	-0.0037	2.3826	-0.0336	-1.39%	1.39%
Benton	2.7971	2.6516	2.6795	-0.0053	2.5137	-0.1658	-6.19%	6.19%
Chelan	2.4827	2.4863	2.6192	0.0001	2.4899	-0.1293	-4.93%	4.93%
Clallam	2.5374	2.4007	2.3066	-0.0055	2.2714	-0.0353	-1.53%	1.53%
Clark	2.7625	2.6625	2.6900	-0.0037	2.5661	-0.1239	-4.61%	4.61%
Columbia	2.5254	2.4368	2.3628	-0.0036	2.3513	-0.0115	-0.49%	0.49%
Cowlitz	2.6619	2.5588	2.5531	-0.0039	2.4597	-0.0934	-3.66%	3.66%
Douglas	2.7591	2.6769	2.7554	-0.0030	2.5971	-0.1583	-5.74%	5.74%
Ferry	2.8567	2.6978	2.4938	-0.0057	2.5477	0.0539	2.16%	2.16%
Franklin	2.8817	3.034	3.2637	0.0052	3.1943	-0.0693	-2.12%	2.12%
Garfield	2.5955	2.3948	2.3911	-0.0080	2.2096	-0.1815	-7.59%	7.59%
Grant	2.7986	2.7407	2.9204	-0.0021	2.6840	-0.2364	-8.09%	8.09%
Grays Harbor	2.5966	2.4813	2.4826	-0.0045	2.3711	-0.1115	-4.49%	4.49%
Island	2.6706	2.6149	2.5223	-0.0021	2.5604	0.0381	1.51%	1.51%
Jefferson	2.4537	2.3089	2.2122	-0.0061	2.1726	-0.0395	-1.79%	1.79%
King	2.4868	2.3982	2.3905	-0.0036	2.3128	-0.0777	-3.25%	3.25%
Kitsap	2.682	2.6469	2.6007	-0.0013	2.6123	0.0115	0.44%	0.44%
Kittitas	2.3976	2.3251	2.3314	-0.0031	2.2548	-0.0766	-3.29%	3.29%
Klickitat	2.7211	2.6409	2.5361	-0.0030	2.5631	0.0270	1.06%	1.06%
Lewis	2.6732	2.5997	2.5690	-0.0028	2.5282	-0.0408	-1.59%	1.59%
Lincoln	2.5726	2.4308	2.4233	-0.0057	2.2968	-0.1265	-5.22%	5.22%
Mason	2.5458	2.5162	2.4891	-0.0012	2.4869	-0.0022	-0.09%	0.09%
Okanogan	2.6674	2.5877	2.5762	-0.0030	2.5104	-0.0658	-2.56%	2.56%
Pacific	2.4465	2.3499	2.2711	-0.0040	2.2571	-0.0140	-0.62%	0.62%
Pend Oreille	2.8088	2.6029	2.5074	-0.0076	2.4121	-0.0953	-3.80%	3.80%
Pierce	2.6586	2.6231	2.6047	-0.0013	2.5881	-0.0166	-0.64%	0.64%
San Juan	2.2946	2.2489	2.1587	-0.0020	2.2041	0.0454	2.10%	2.10%
Skagit	2.5656	2.5495	2.6032	-0.0006	2.5335	-0.0697	-2.68%	2.68%
Skamania	2.7896	2.6921	2.6120	-0.0036	2.5980	-0.0140	-0.54%	0.54%
Snohomish	2.7606	2.67935	2.6547	-0.0030	2.6005	-0.0542	-2.04%	2.04%
Spokane	2.5789	2.4747	2.4646	-0.0041	2.3747	-0.0899	-3.65%	3.65%
Stevens	2.907	2.7318	2.6439	-0.0062	2.5672	-0.0768	-2.90%	2.90%
Thurston	2.6441	2.553	2.4987	-0.0035	2.4650	-0.0337	-1.35%	1.35%
Wahkiakum	2.7724	2.4762	2.4243	-0.0112	2.2116	-0.2127	-8.77%	8.77%
Walla Walla	2.5411	2.4955	2.5388	-0.0018	2.4507	-0.0880	-3.47%	3.47%
Whatcom	2.5902	2.5324	2.5113	-0.0023	2.4759	-0.0354	-1.41%	1.41%
Whitman	2.4668	2.3868	2.3115	-0.0033	2.3094	-0.0021	-0.09%	0.09%
Yakima	2.7711	2.8039	2.9576	0.0012	2.8371	-0.1205	-4.08%	4.08%

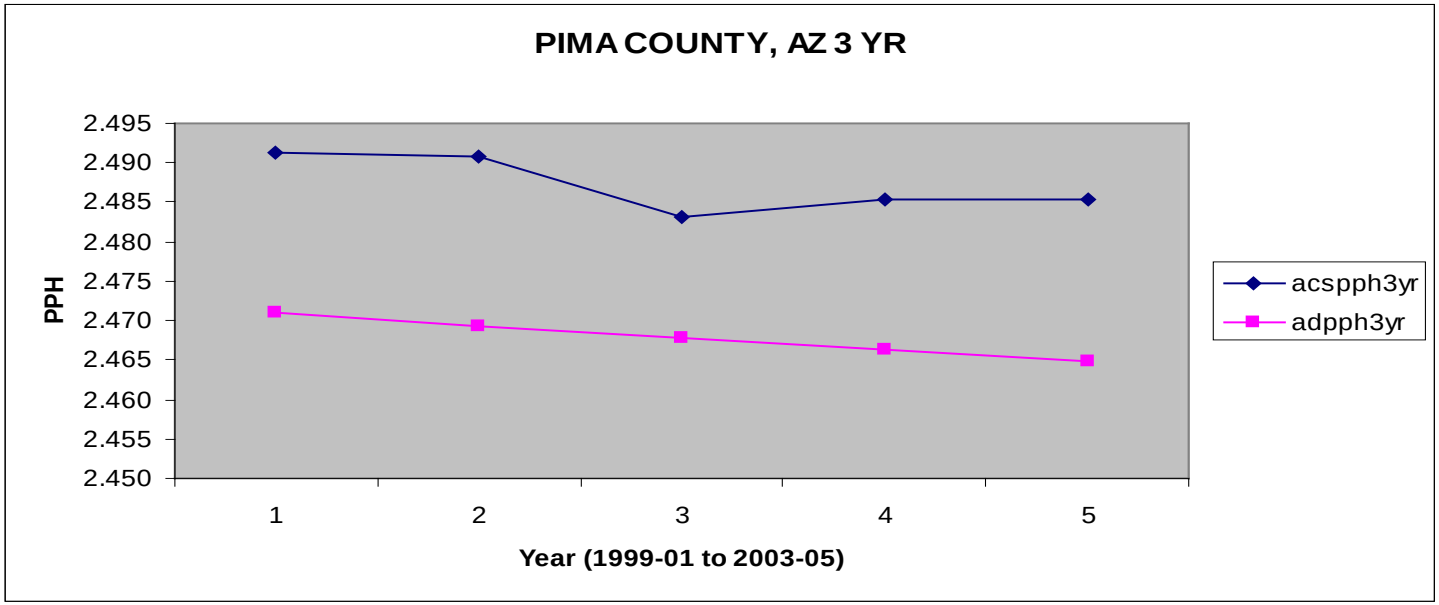
DATA

Accuracy Test of the Geometric Method of Estimating PPH Values for Counties: Washington State 2000

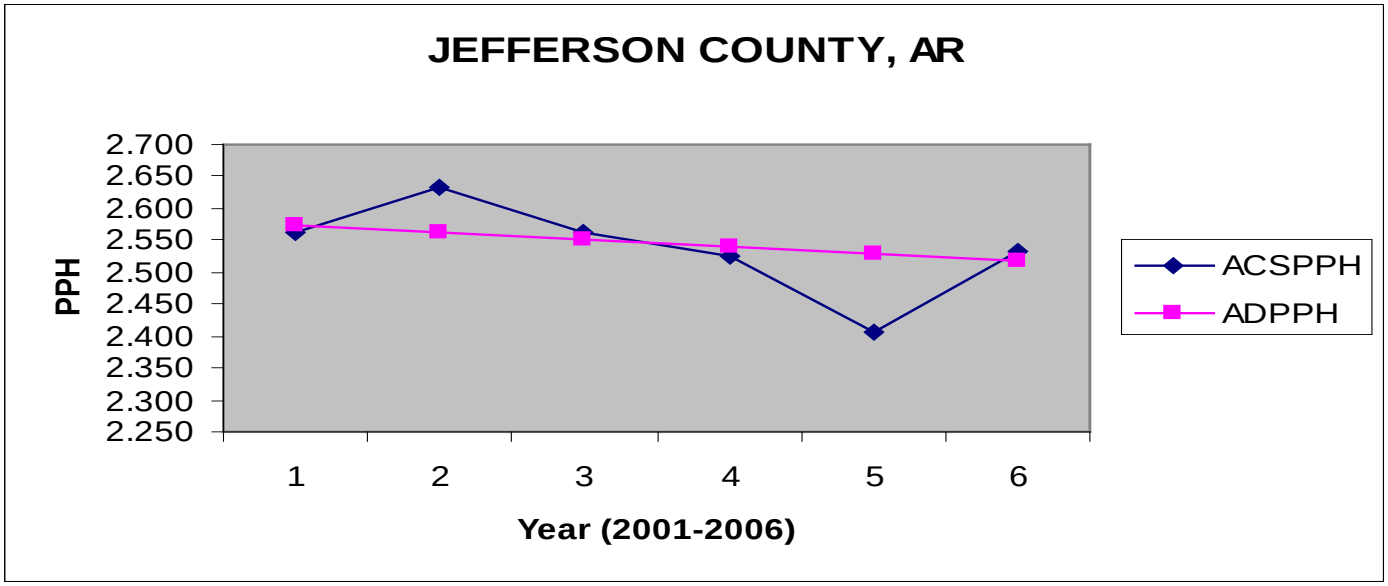
COUNTY LEVEL SUMMARY STATISTICS	
Mean Error	-0.0680
MAPE	2.97%
MALPE	-2.60%
N ABS % ERROR >5	6

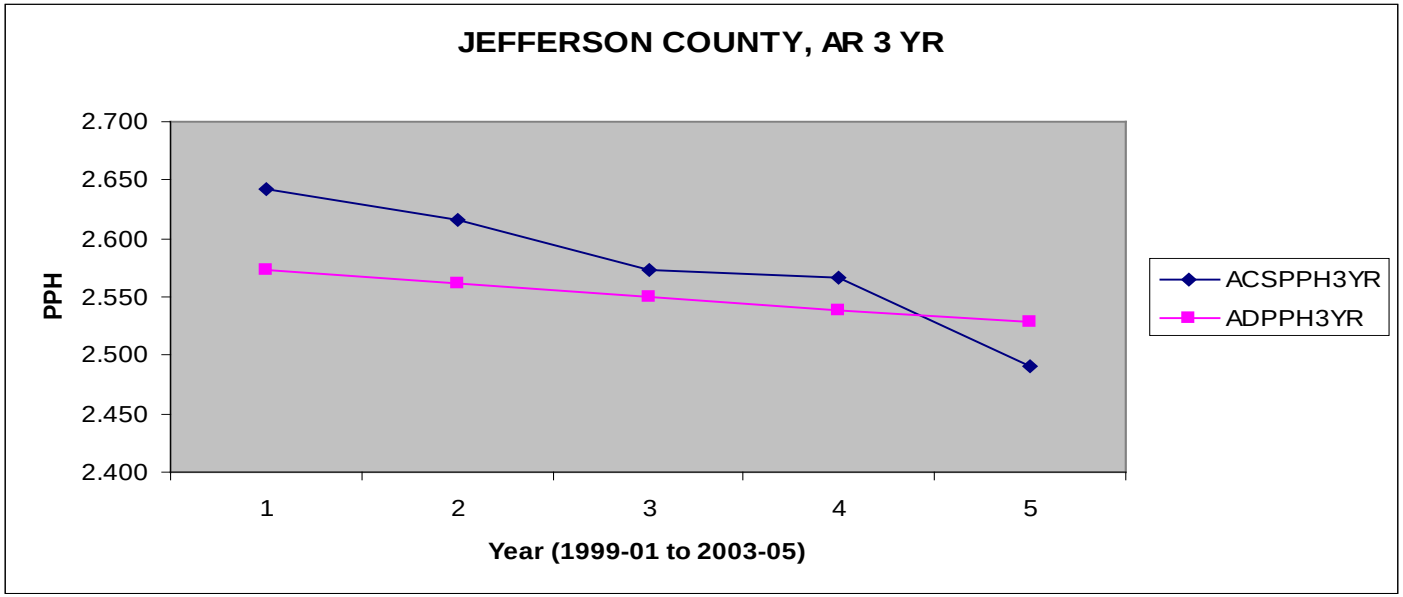
RESULTS


PIMA COUNTY, AZ 2001-2006
PPH
2.520
2.510
2.500
2.490
2.480
2.470
2.460
2.450
2.440
2.430
1
2
3
4
5
6
Year (2001 to 2006)
ACSPPH
ADPPH

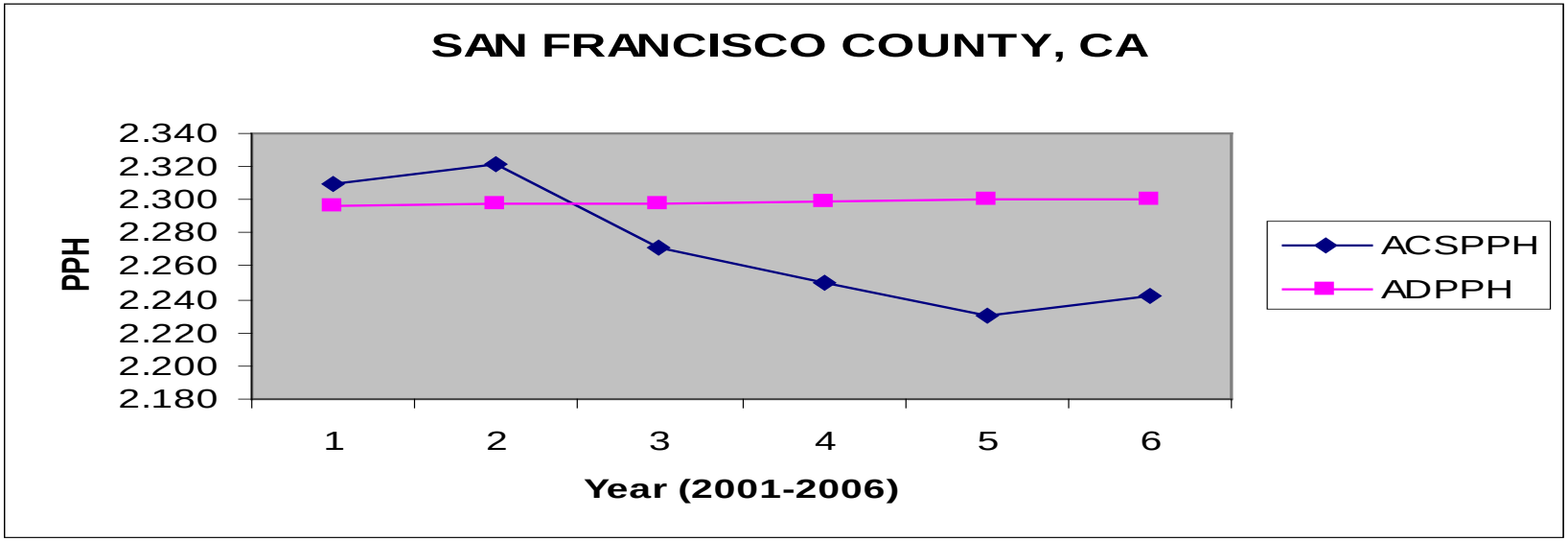

PIMA COUNTY, AZ 3 YR
PPH
2.495
2.490
2.485
2.480
2.475
2.470
2.465
2.460
2.455
2.450
1
2
3
4
5
Year (1999-01 to 2003-05)
acspph3yr
adpph3yr

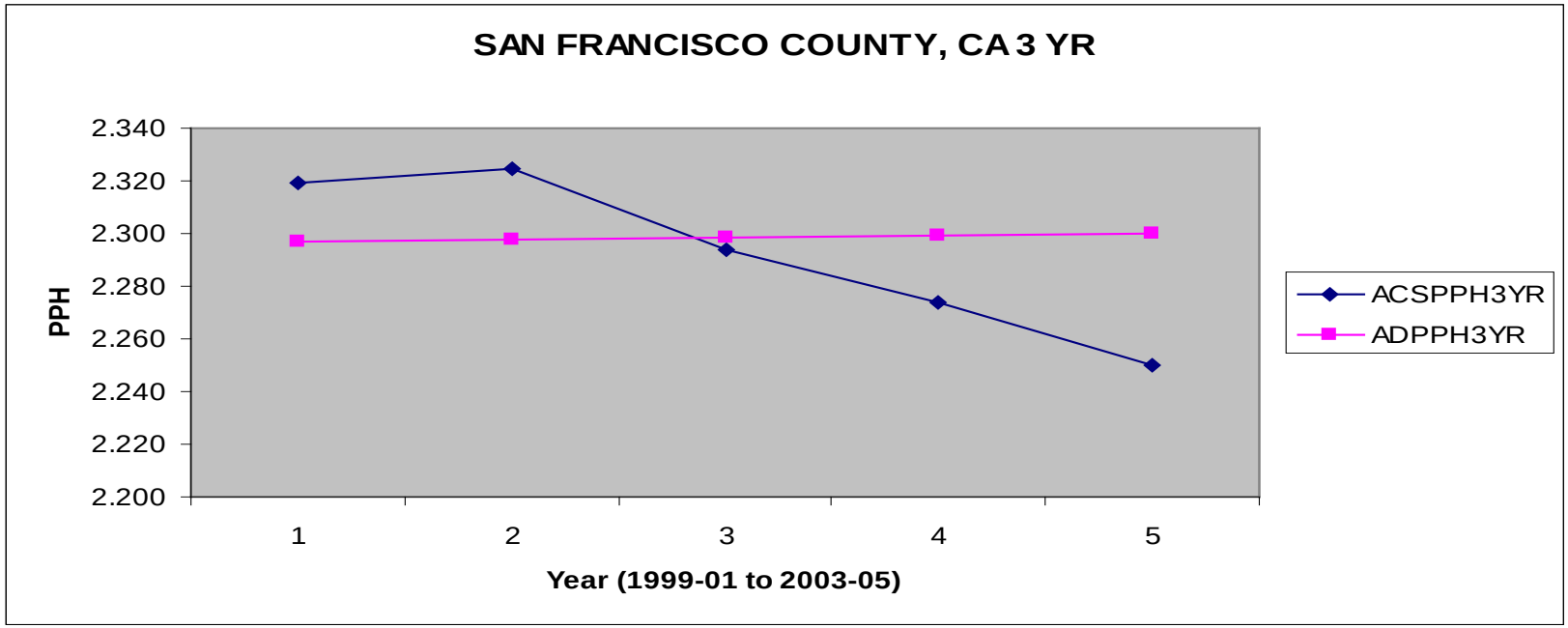
RESULTS


JEFFERSON COUNTY, AR
PPH
2.700
2.650
2.600
2.550
2.500
2.450
2.400
2.350
2.300
2.250
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

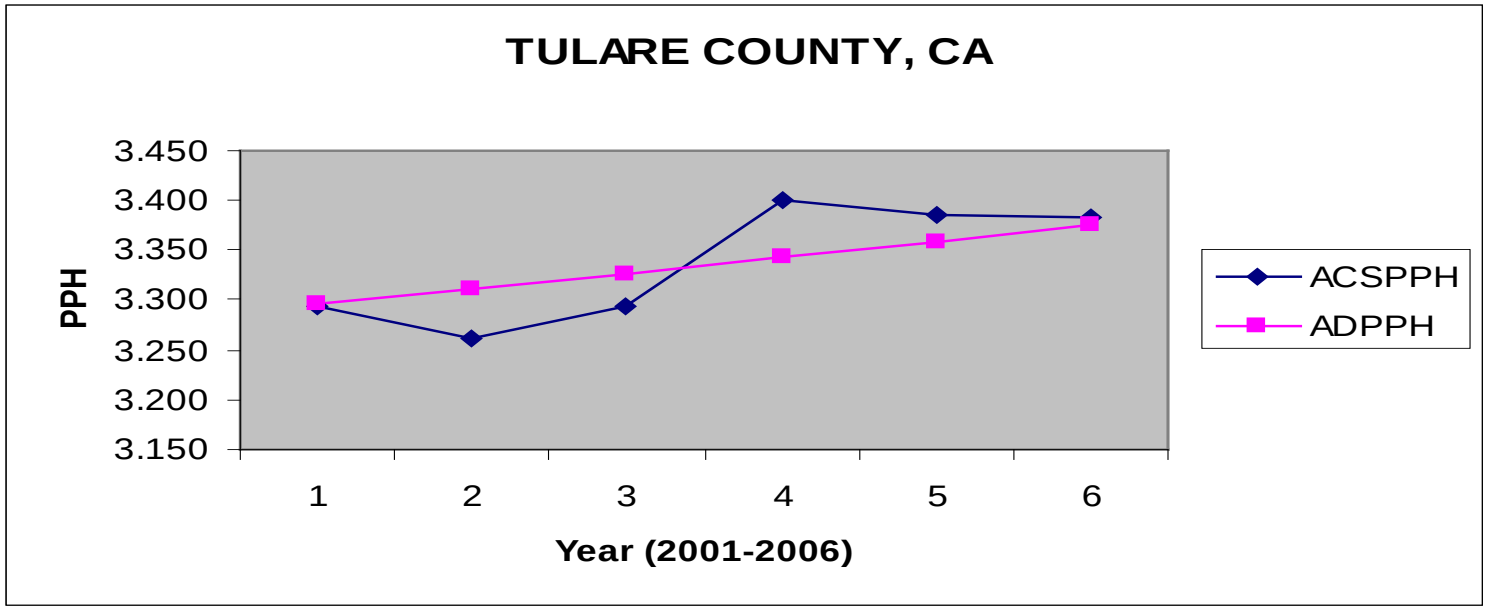

JEFFERSON COUNTY, AR 3 YR
PPH
2.700
2.650
2.600
2.550
2.500
2.450
2.400
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

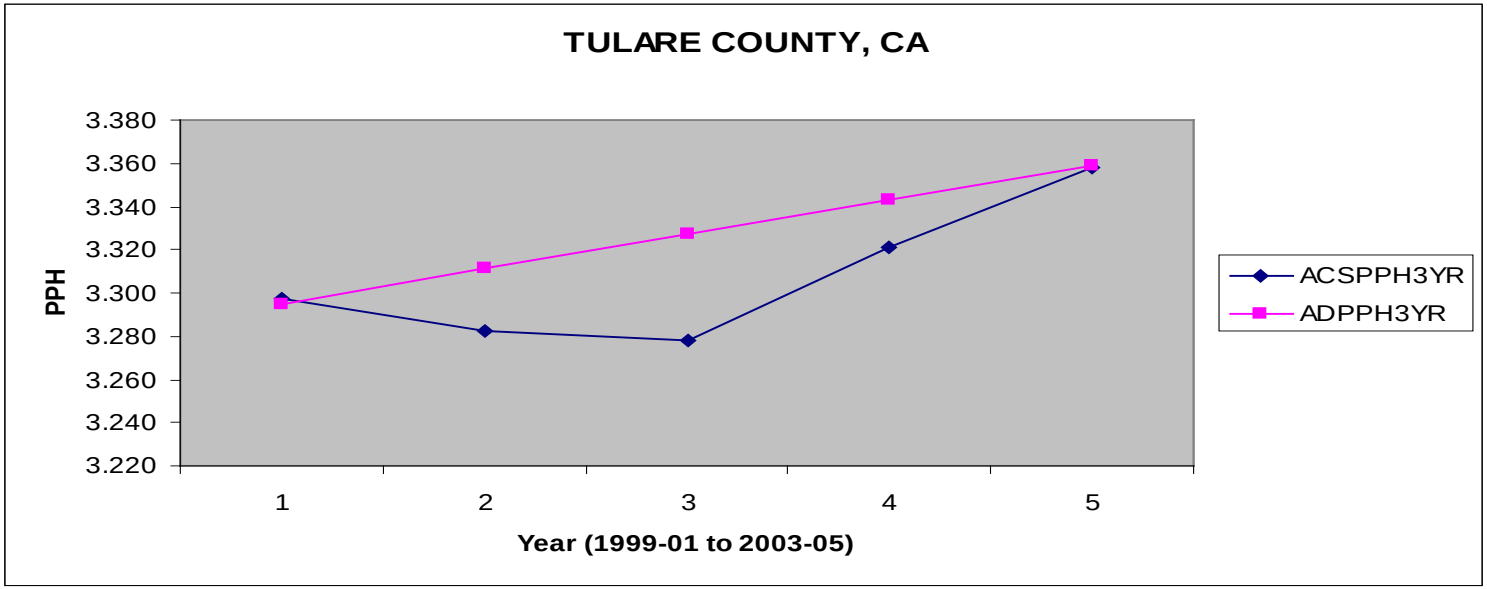
RESULTS


SAN FRANCISCO COUNTY, CA
PPH
2.340
2.320
2.300
2.280
2.260
2.240
2.220
2.200
2.180
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

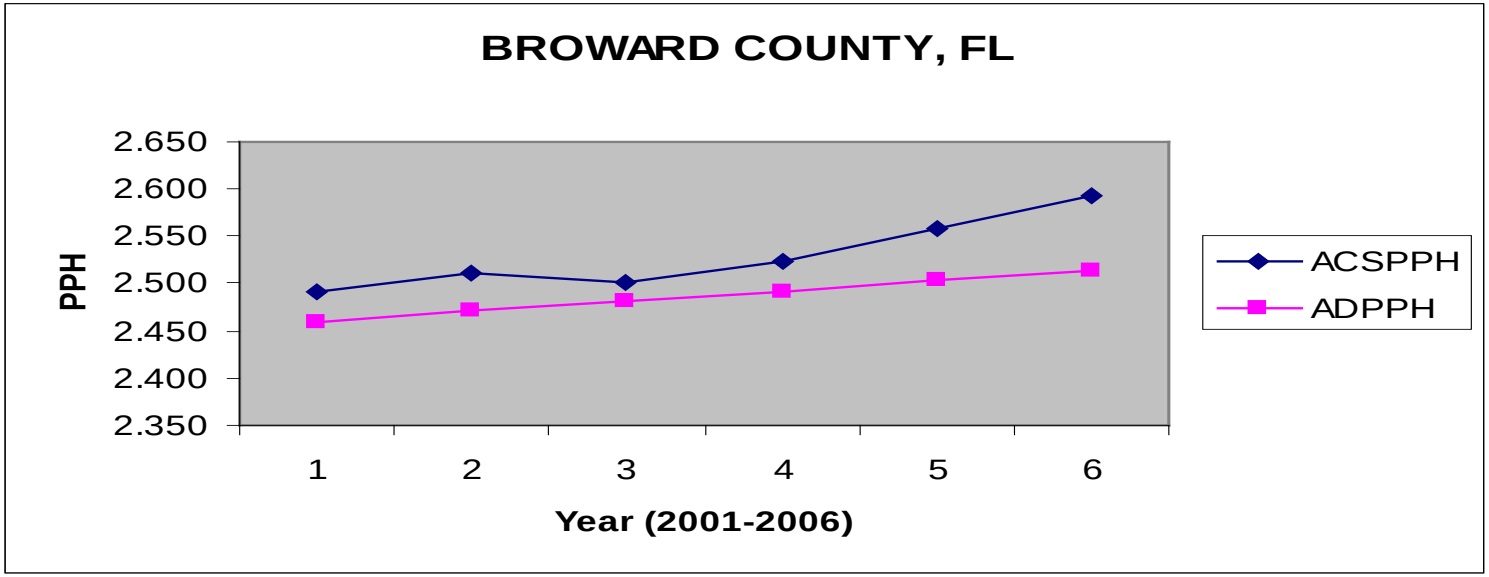

SAN FRANCISCO COUNTY, CA 3 YR
PPH
2.340
2.320
2.300
2.280
2.260
2.240
2.220
2.200
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

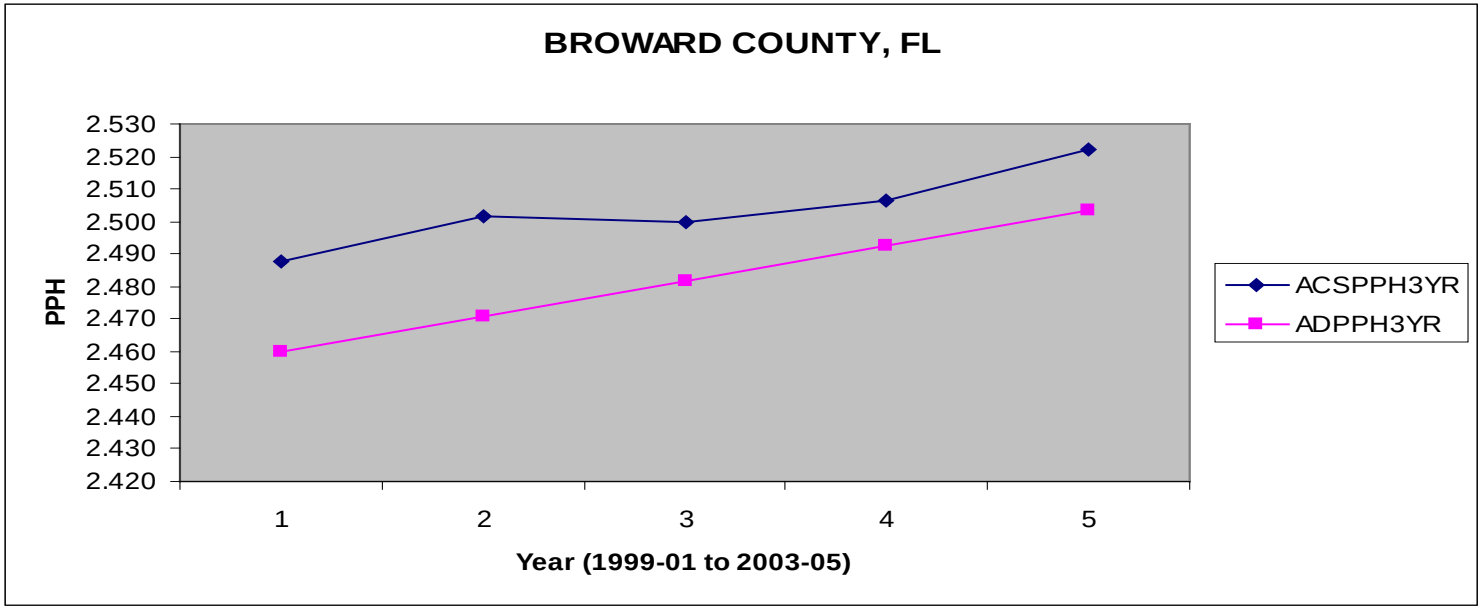
RESULTS


TULARE COUNTY, CA
3.450
3.400
3.350
3.300
3.250
3.200
3.150
PPH
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

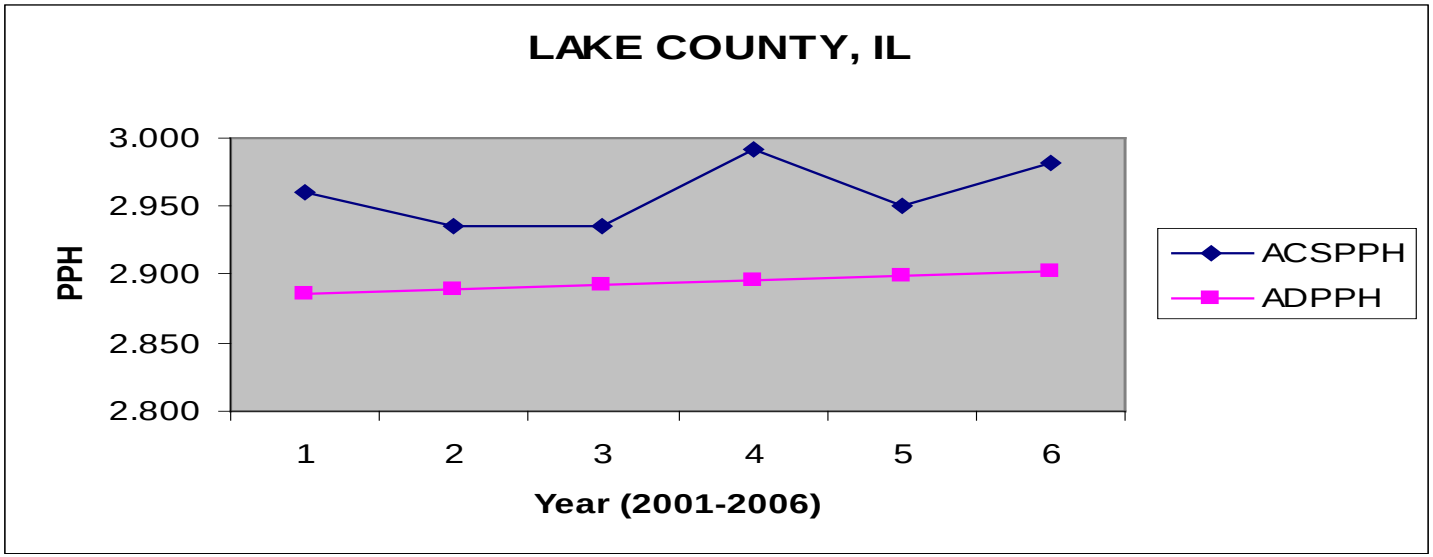

TULARE COUNTY, CA
3.380
3.360
3.340
3.320
3.300
3.280
3.260
3.240
3.220
PPH
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

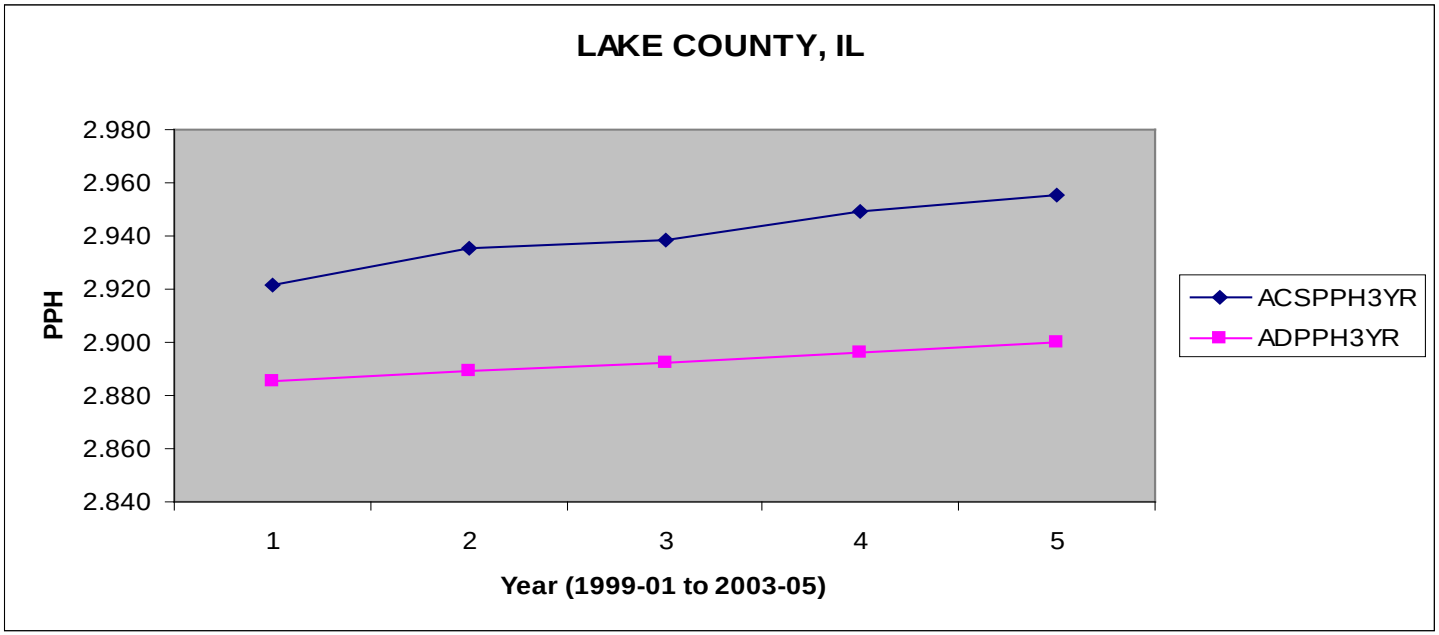
RESULTS


BROWARD COUNTY, FL
PPH
2.650
2.600
2.550
2.500
2.450
2.400
2.350
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

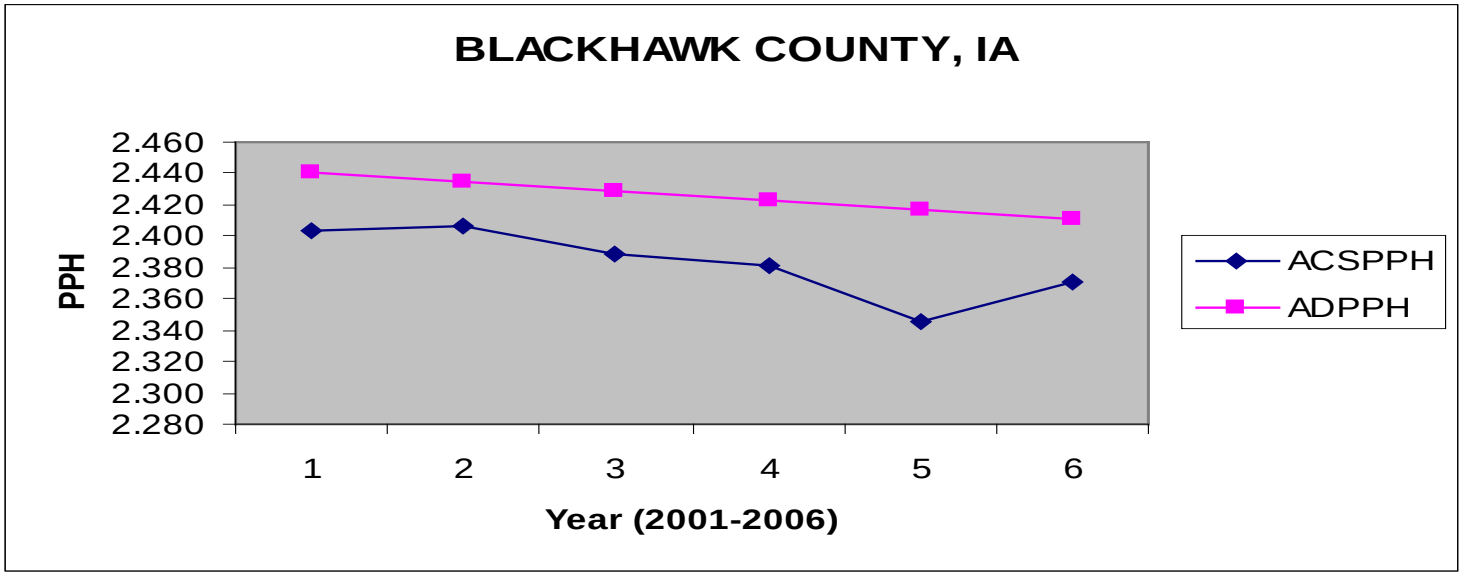

BROWARD COUNTY, FL
PPH
2.530
2.520
2.510
2.500
2.490
2.480
2.470
2.460
2.450
2.440
2.430
2.420
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

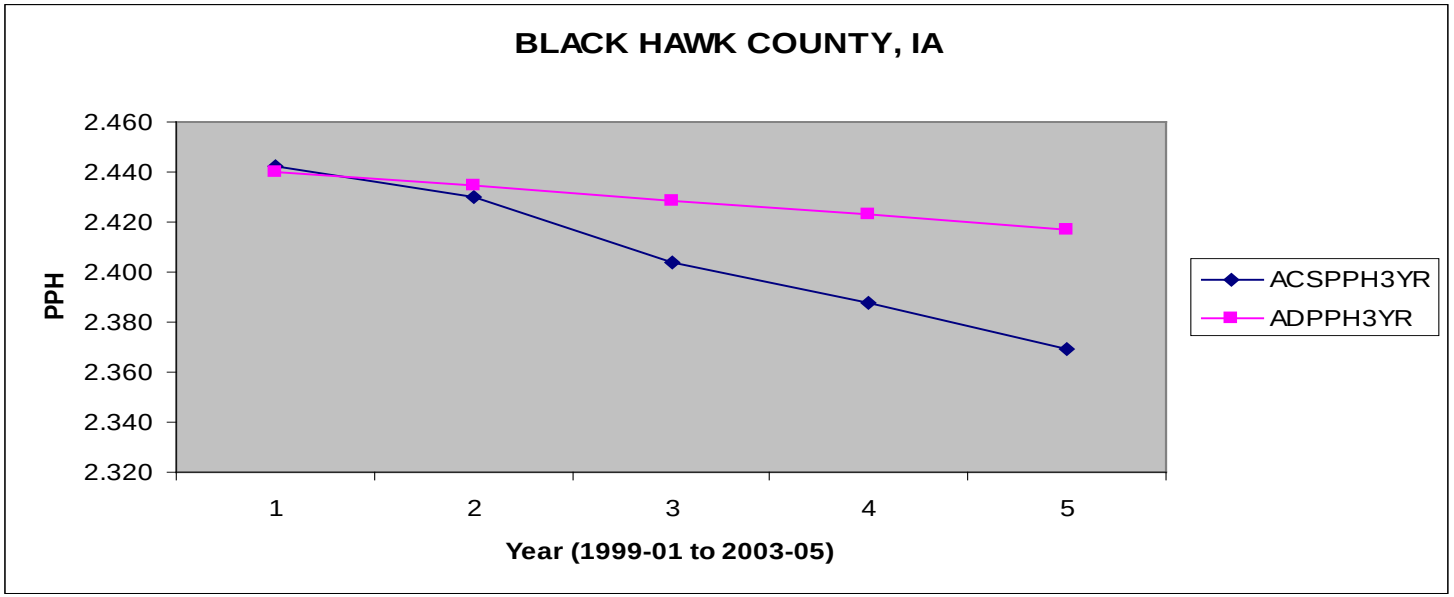
RESULTS


LAKE COUNTY, IL
3.000
2.950
2.900
2.850
2.800
PPH
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

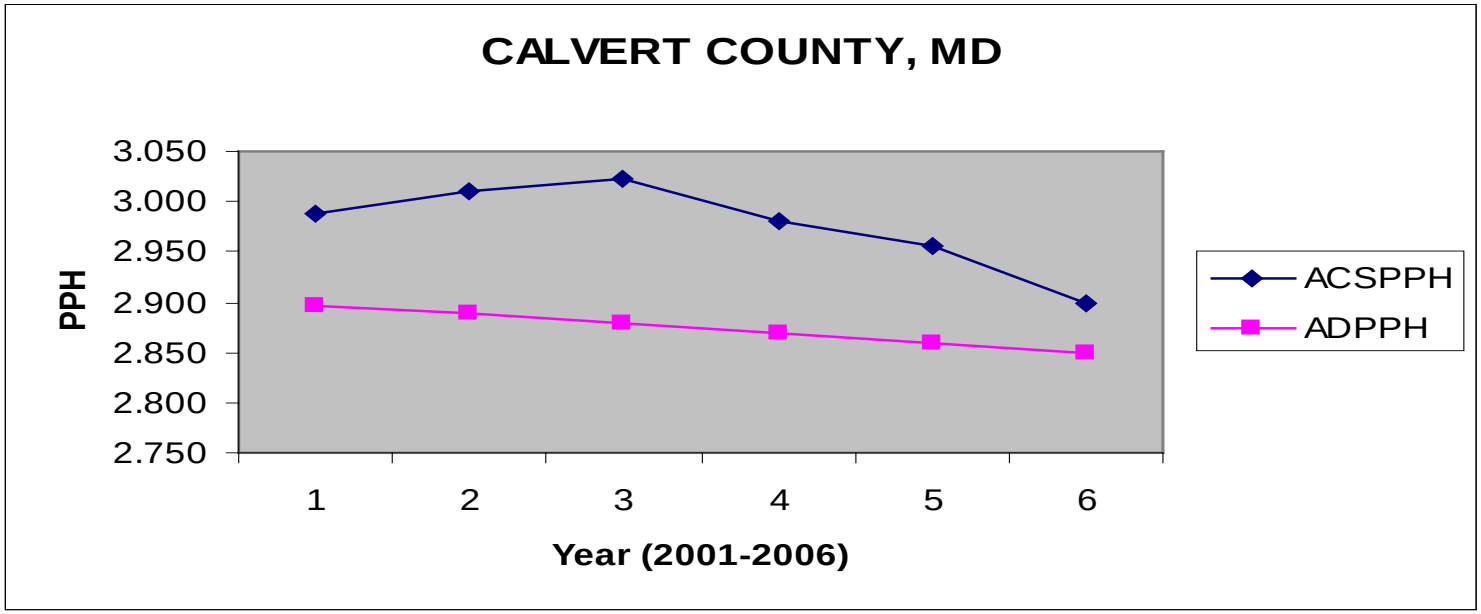

LAKE COUNTY, IL
2.980
2.960
2.940
2.920
2.900
2.880
2.860
2.840
PPH
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

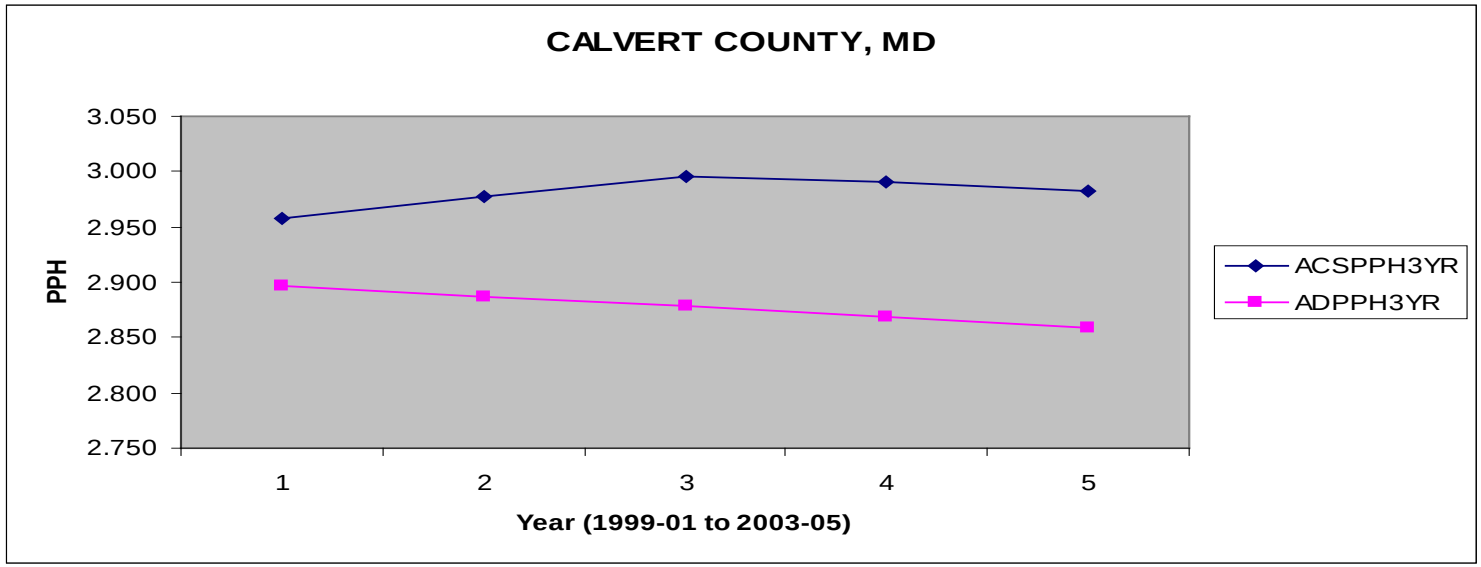
RESULTS


BLACKHAWK COUNTY, IA
PPH
2.460
2.440
2.420
2.400
2.380
2.360
2.340
2.320
2.300
2.280
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

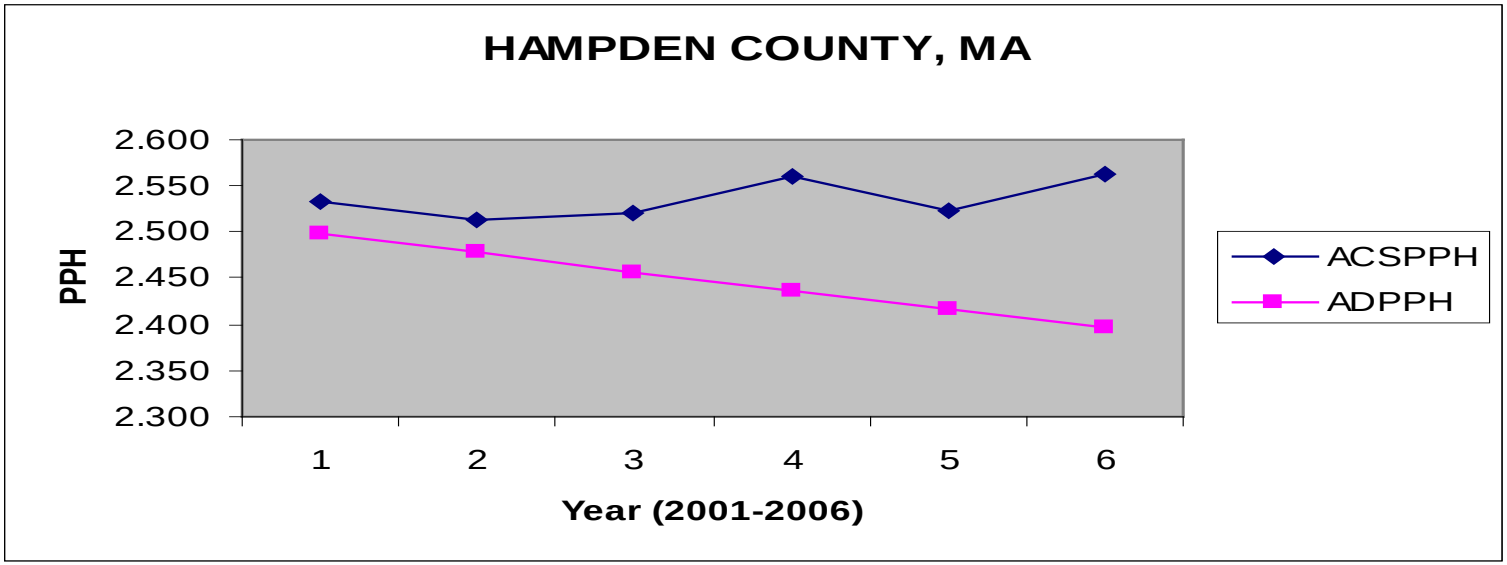

BLACK HAWK COUNTY, IA
PPH
2.460
2.440
2.420
2.400
2.380
2.360
2.340
2.320
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

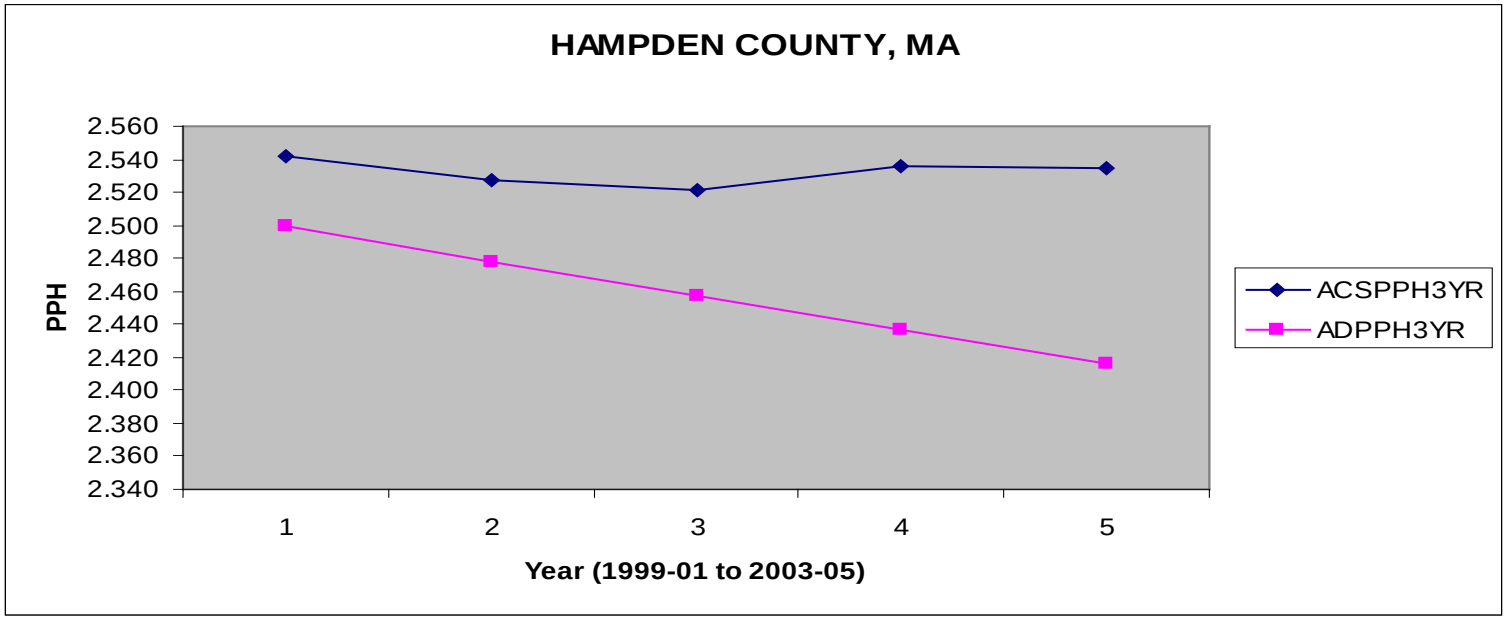
RESULTS


CALVERT COUNTY, MD
3.050
3.000
2.950
2.900
2.850
2.800
2.750
PPH
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

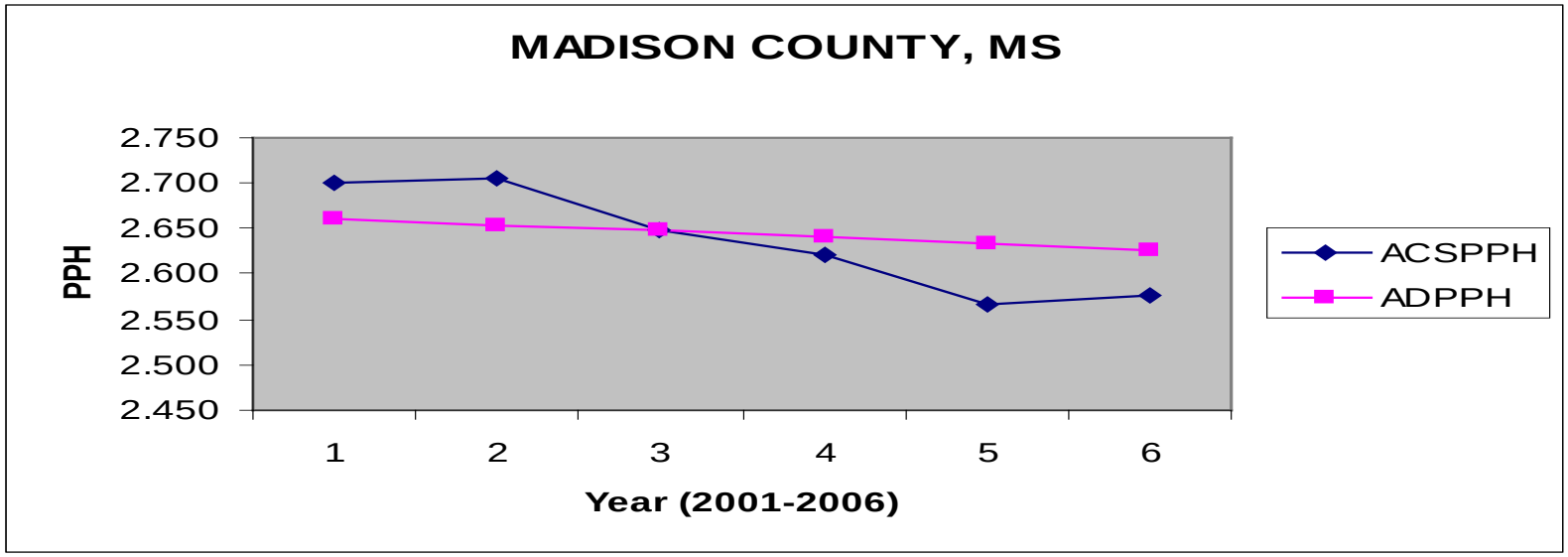

CALVERT COUNTY, MD
3.050
3.000
2.950
2.900
2.850
2.800
2.750
PPH
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

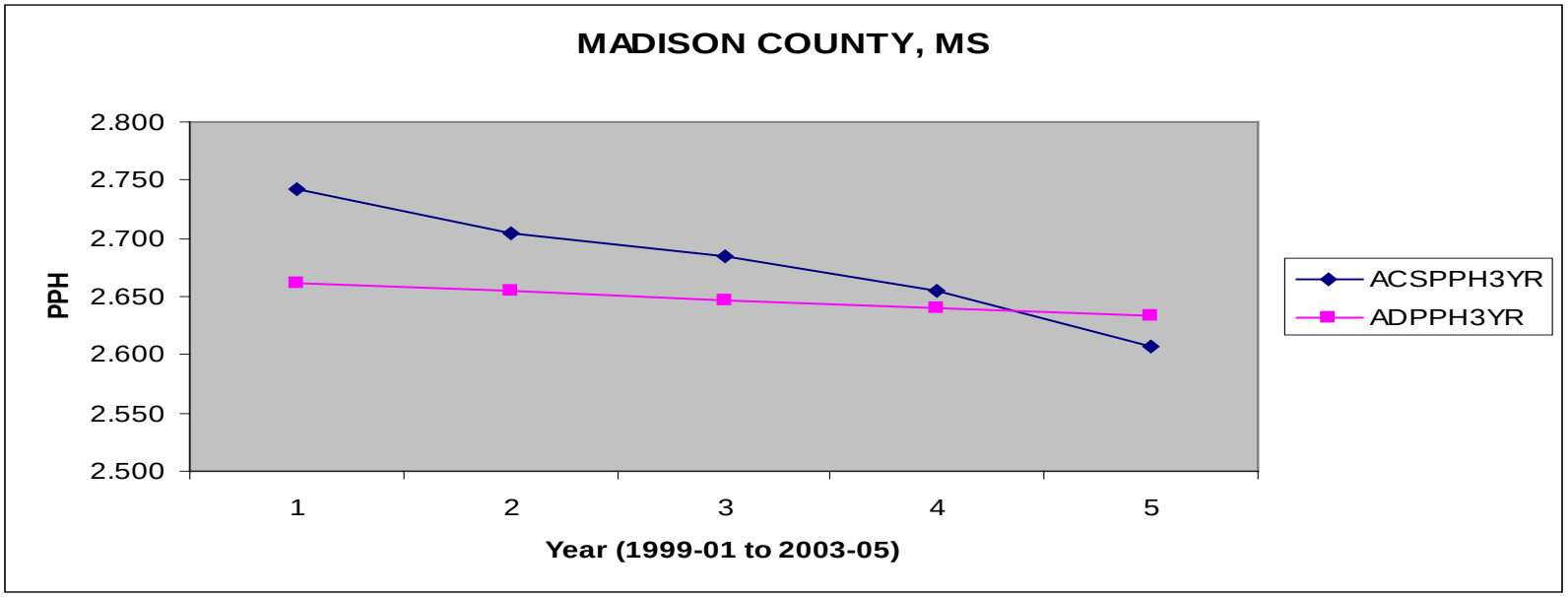
RESULTS


HAMPDEN COUNTY, MA
PPH
2.600
2.550
2.500
2.450
2.400
2.350
2.300
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

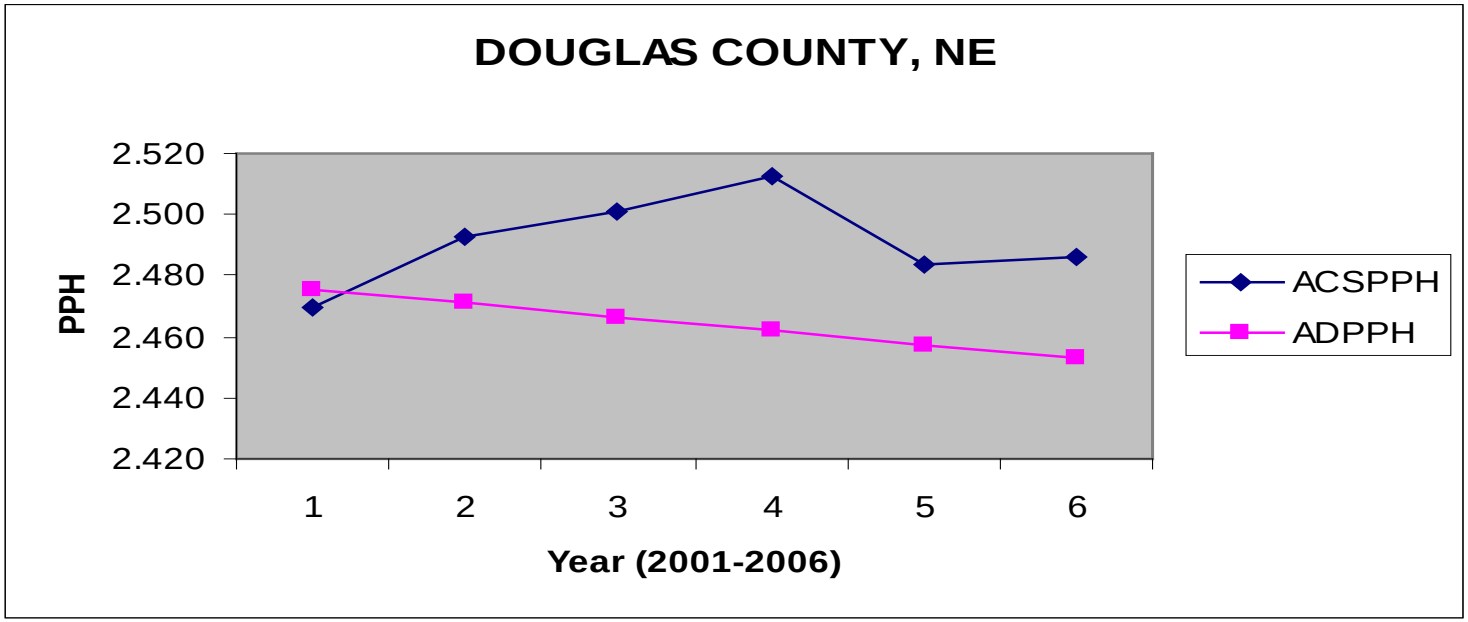

HAMPDEN COUNTY, MA
PPH
2.560
2.540
2.520
2.500
2.480
2.460
2.440
2.420
2.400
2.380
2.360
2.340
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

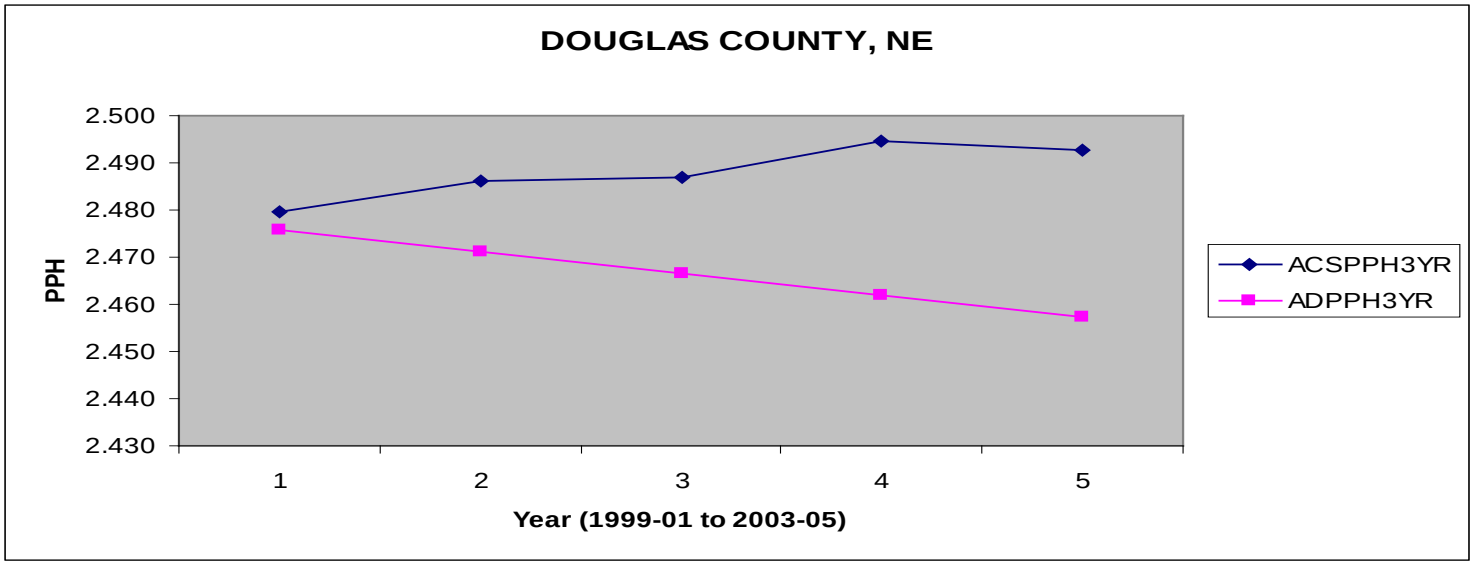
RESULTS


MADISON COUNTY, MS
2.750
2.700
2.650
2.600
2.550
2.500
2.450
PPH
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

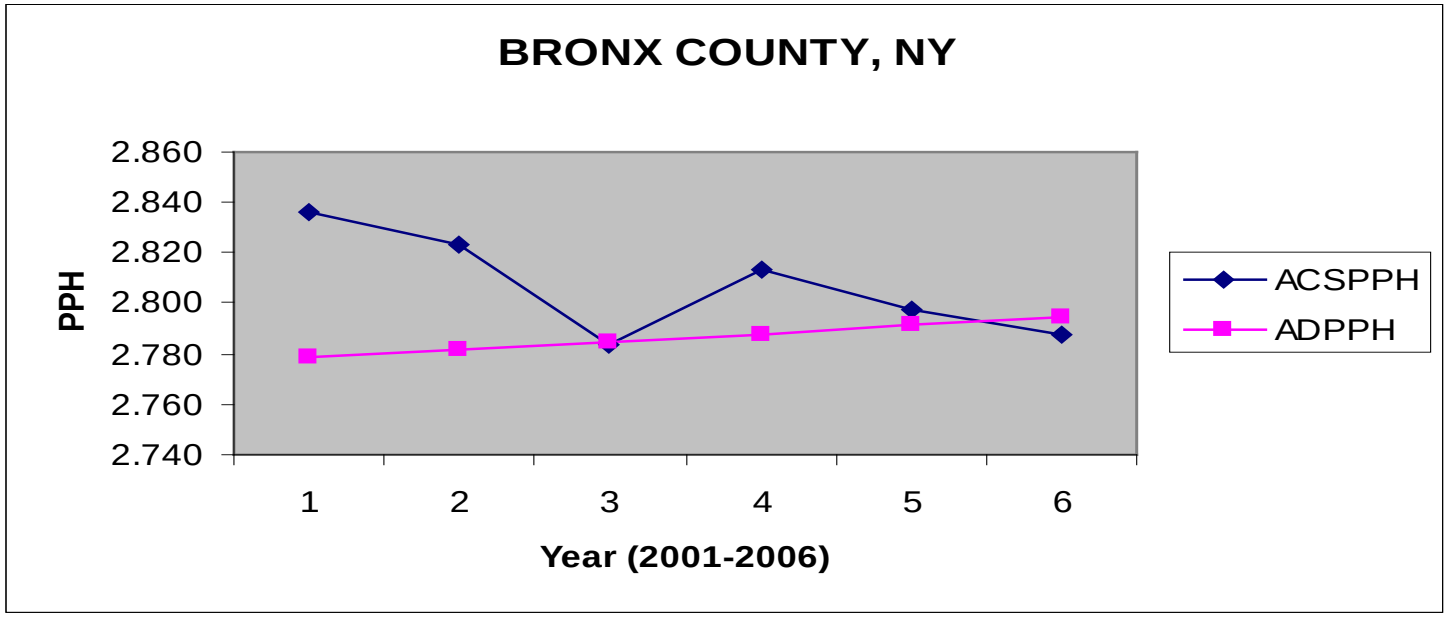

MADISON COUNTY, MS
2.800
2.750
2.700
2.650
2.600
2.550
2.500
PPH
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

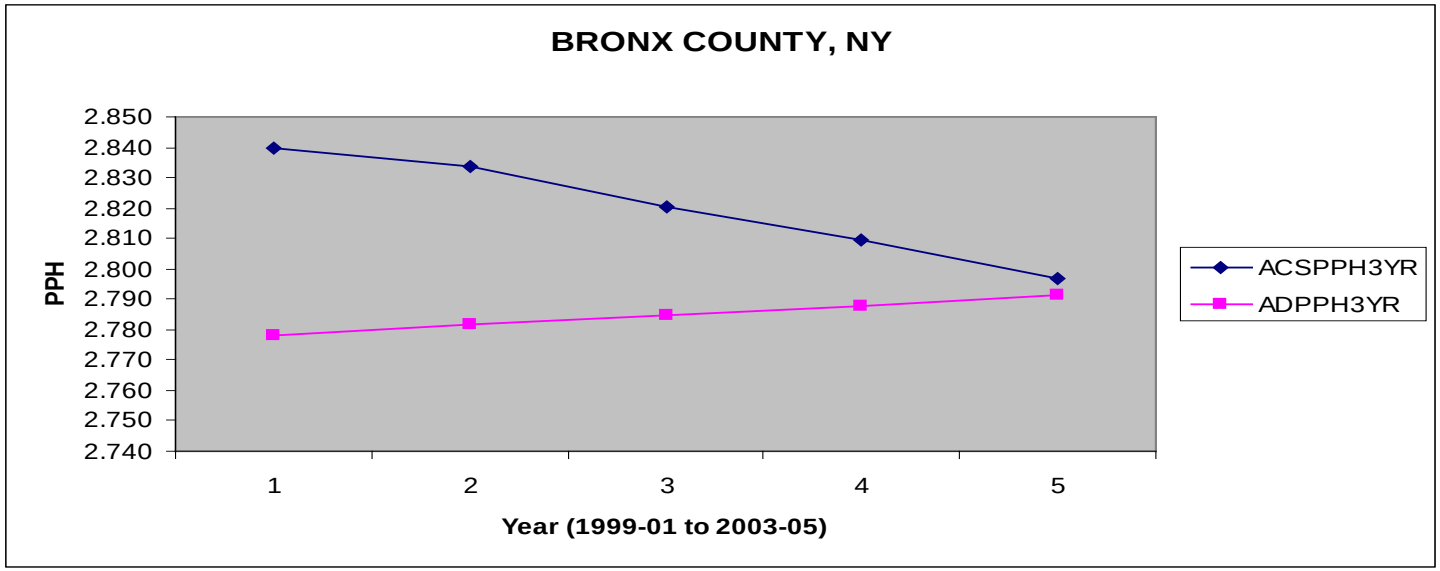
RESULTS


DOUGLAS COUNTY, NE
PPH
2.520
2.500
2.480
2.460
2.440
2.420
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

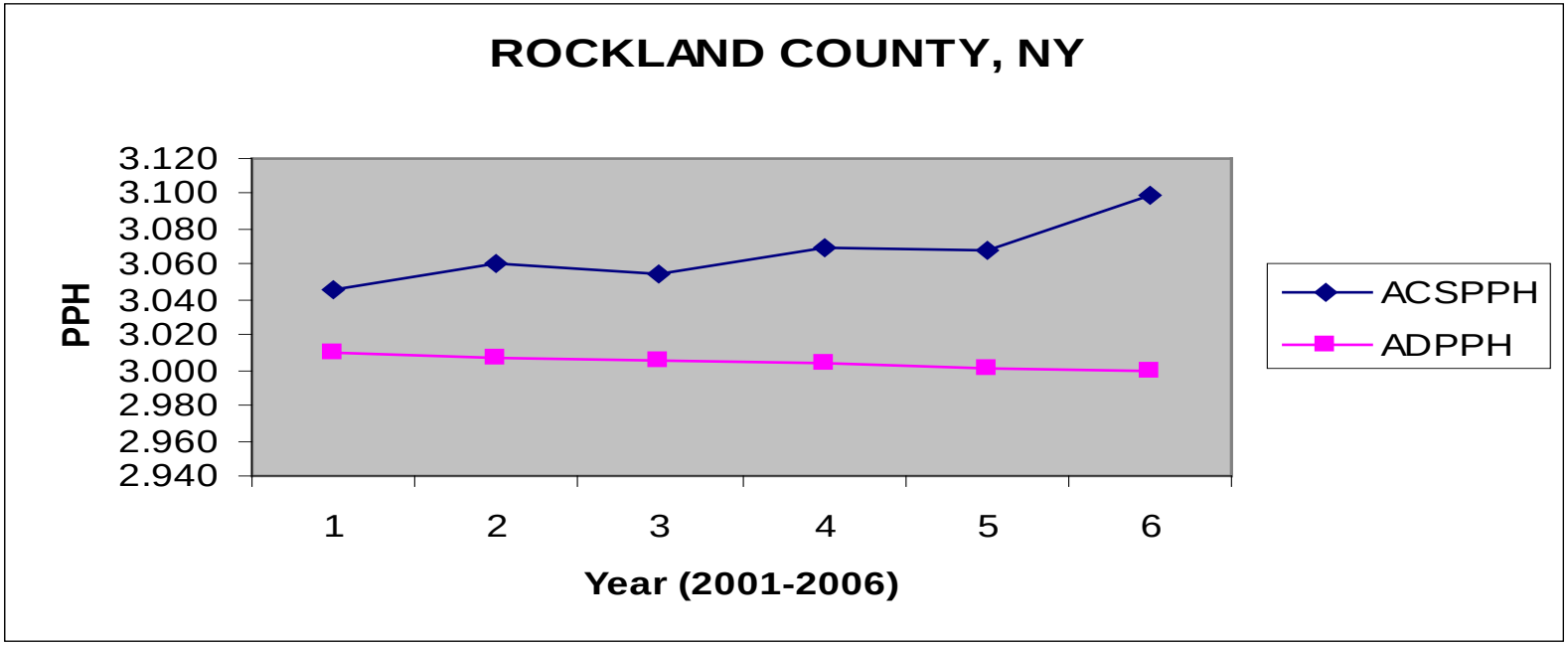

DOUGLAS COUNTY, NE
PPH
2.500
2.490
2.480
2.470
2.460
2.450
2.440
2.430
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

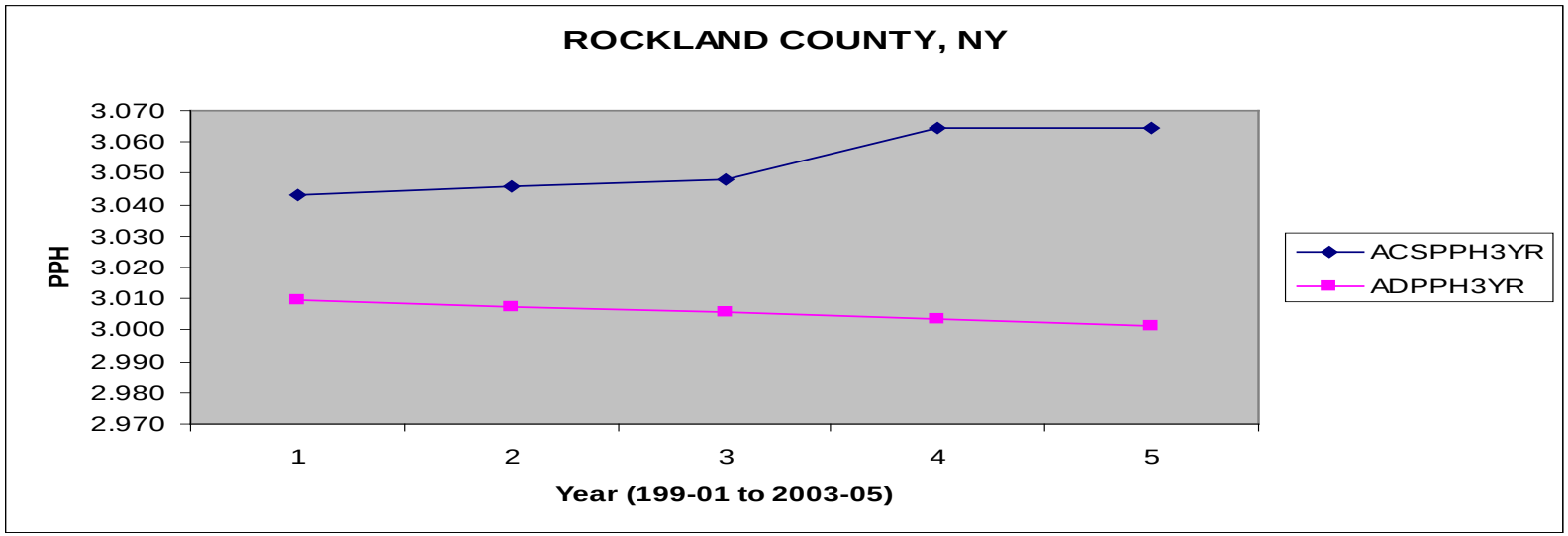
RESULTS


BRONX COUNTY, NY
PPH
2.860
2.840
2.820
2.800
2.780
2.760
2.740
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

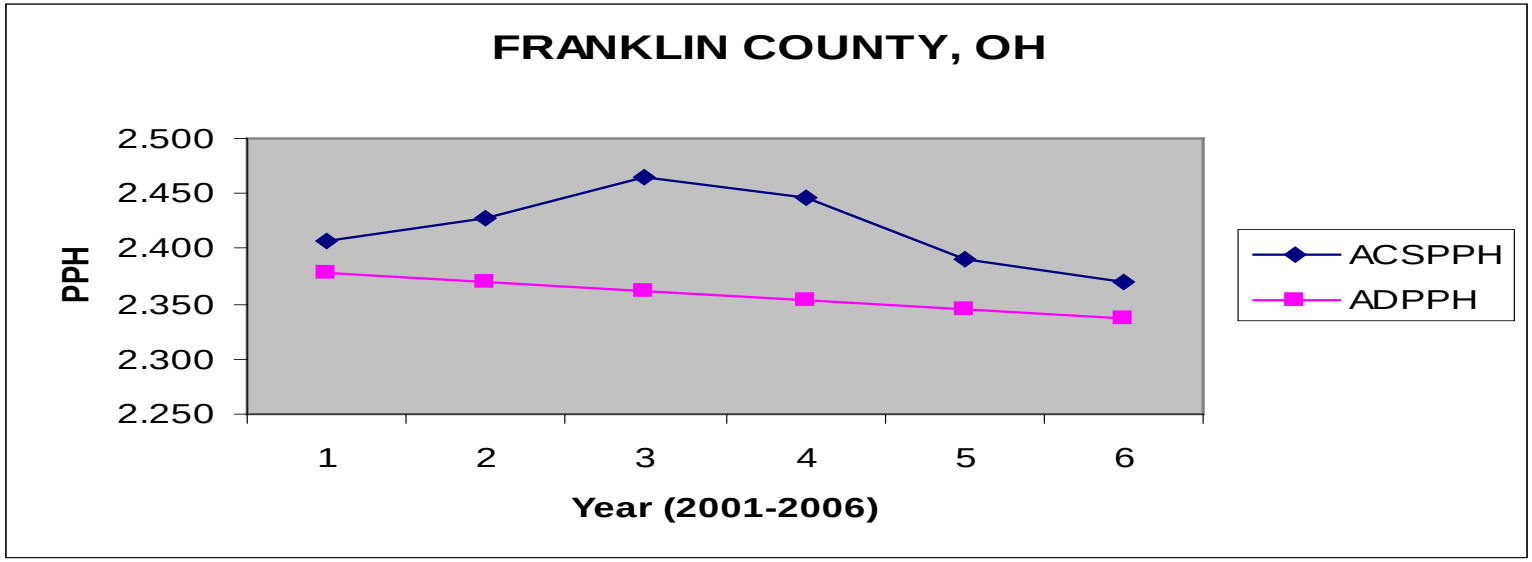

BRONX COUNTY, NY
PPH
2.850
2.840
2.830
2.820
2.810
2.800
2.790
2.780
2.770
2.760
2.750
2.740
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

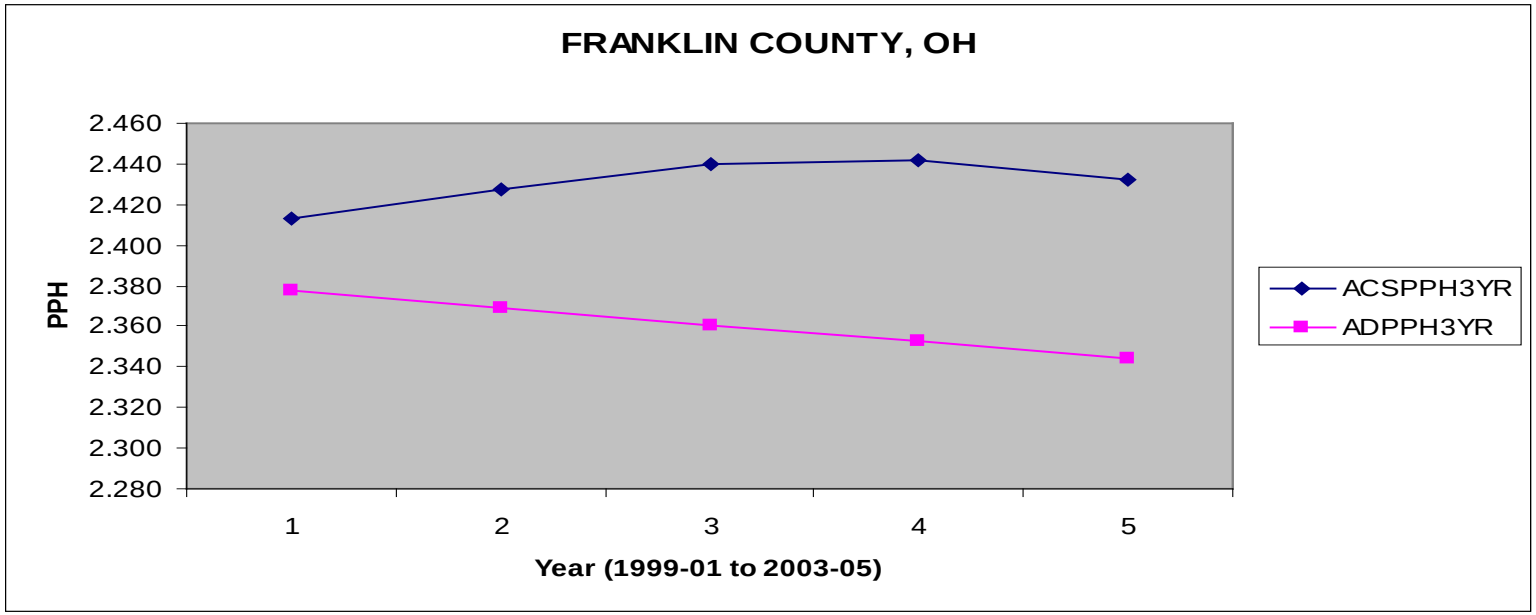
RESULTS


ROCKLAND COUNTY, NY
PPH
3.120
3.100
3.080
3.060
3.040
3.020
3.000
2.980
2.960
2.940
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

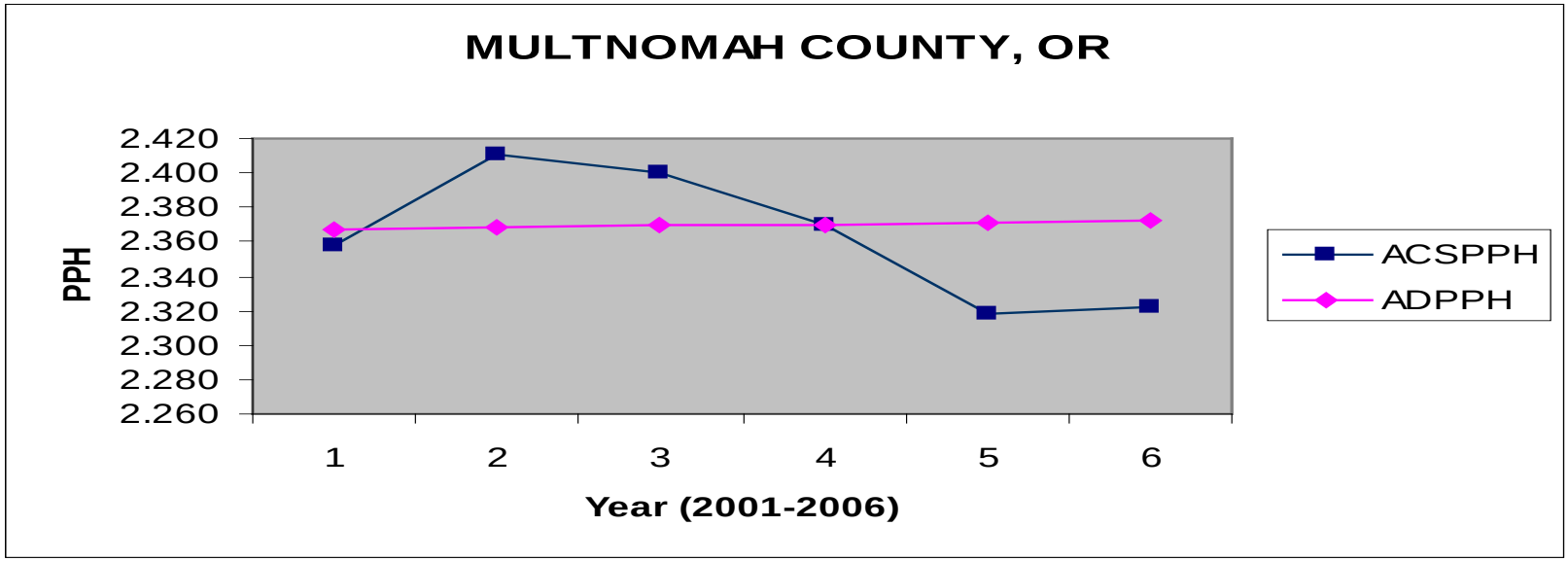

ROCKLAND COUNTY, NY
PPH
3.070
3.060
3.050
3.040
3.030
3.020
3.010
3.000
2.990
2.980
2.970
1
2
3
4
5
Year (199-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

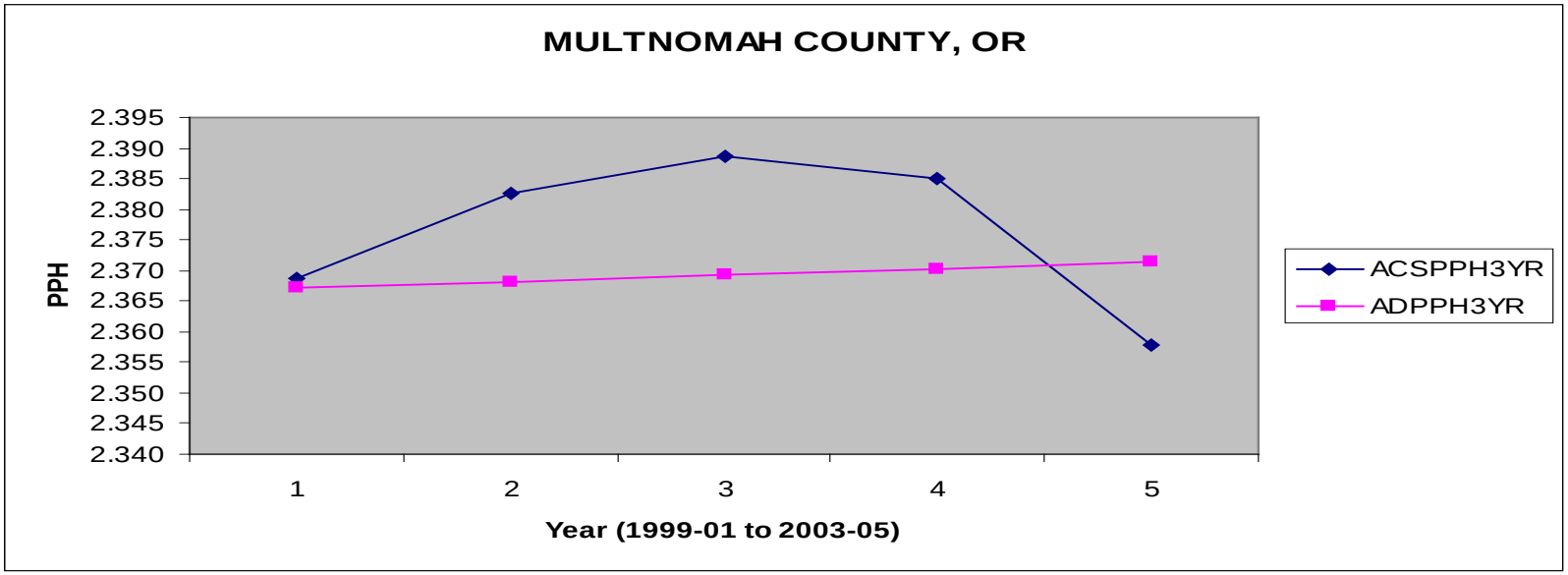
RESULTS


FRANKLIN COUNTY, OH
PPH
2.500
2.450
2.400
2.350
2.300
2.250
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

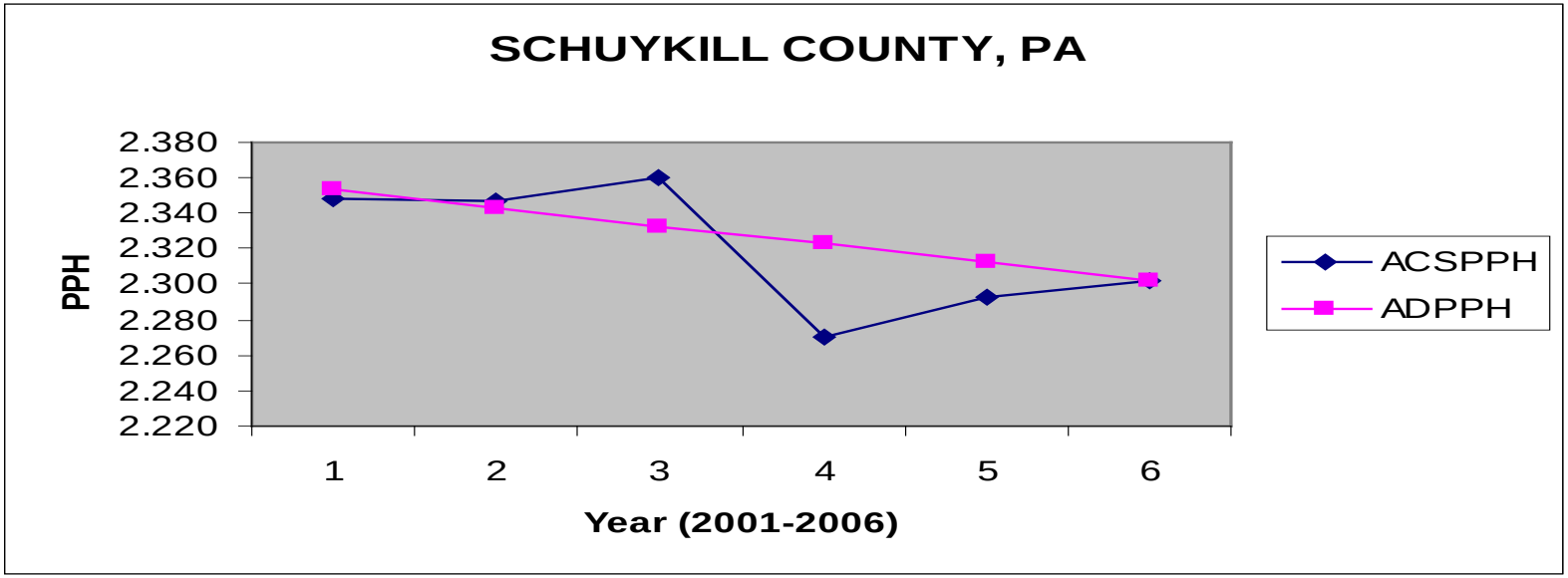

FRANKLIN COUNTY, OH
PPH
2.460
2.440
2.420
2.400
2.380
2.360
2.340
2.320
2.300
2.280
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

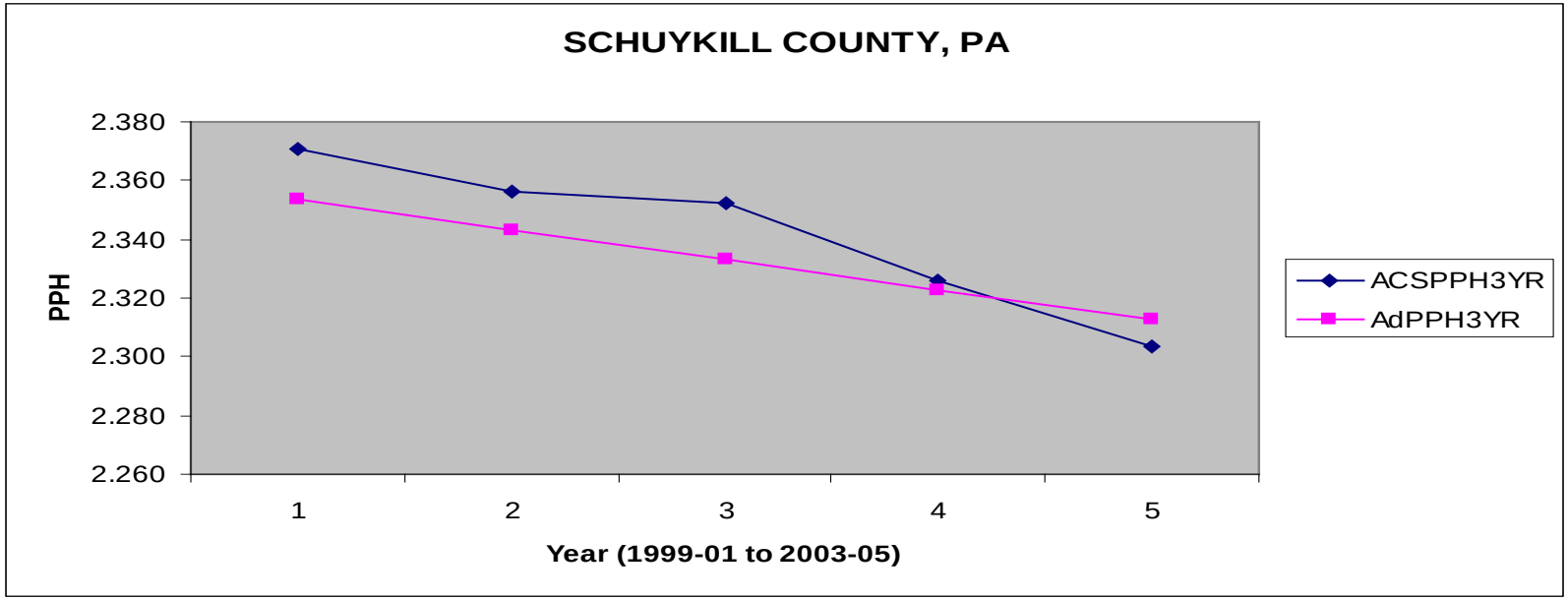
RESULTS


MULTNOMAH COUNTY, OR
PPH
2.420
2.400
2.380
2.360
2.340
2.320
2.300
2.280
2.260
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

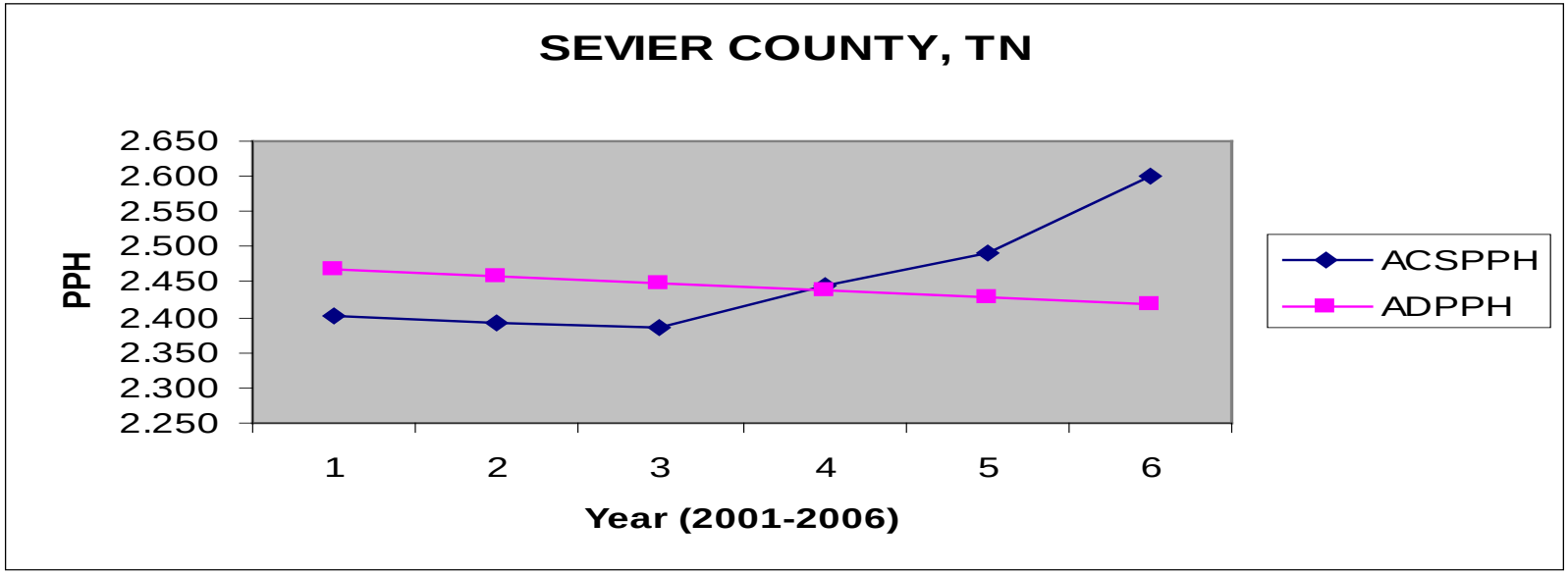

MULTNOMAH COUNTY, OR
PPH
2.395
2.390
2.385
2.380
2.375
2.370
2.365
2.360
2.355
2.350
2.345
2.340
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

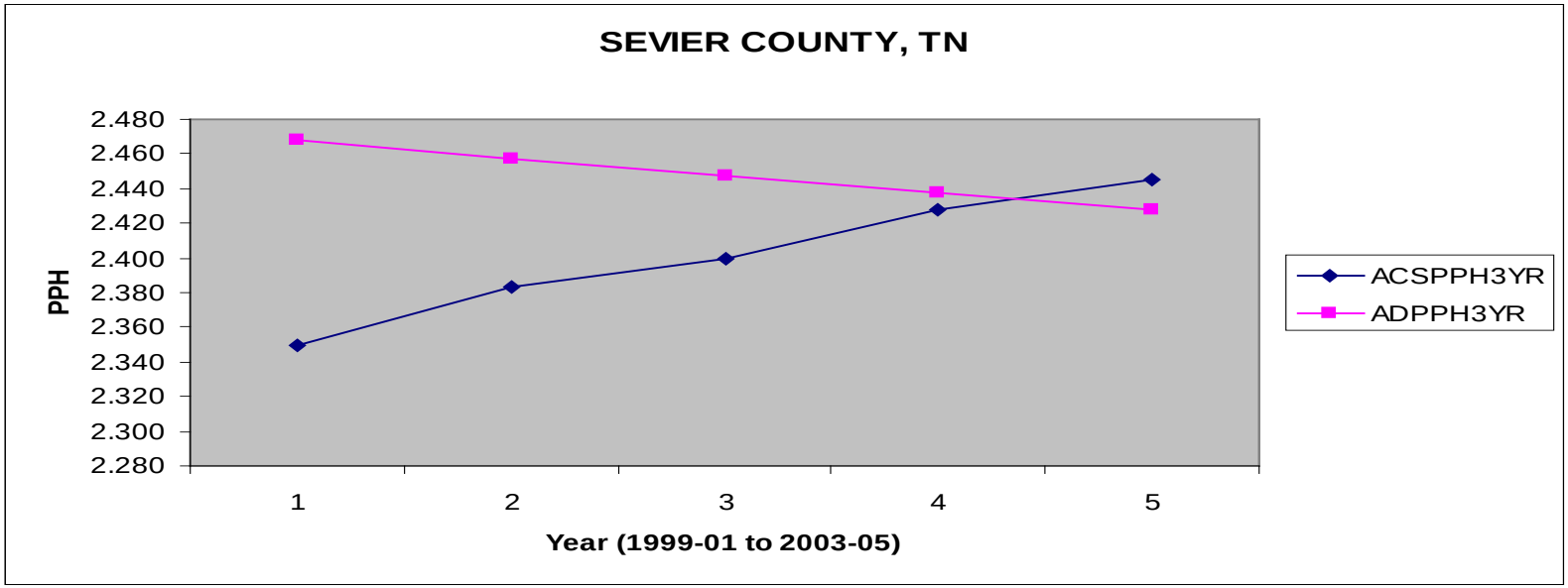
RESULTS


SCHUYKILL COUNTY, PA
PPH
2.380
2.360
2.340
2.320
2.300
2.280
2.260
2.240
2.220
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH

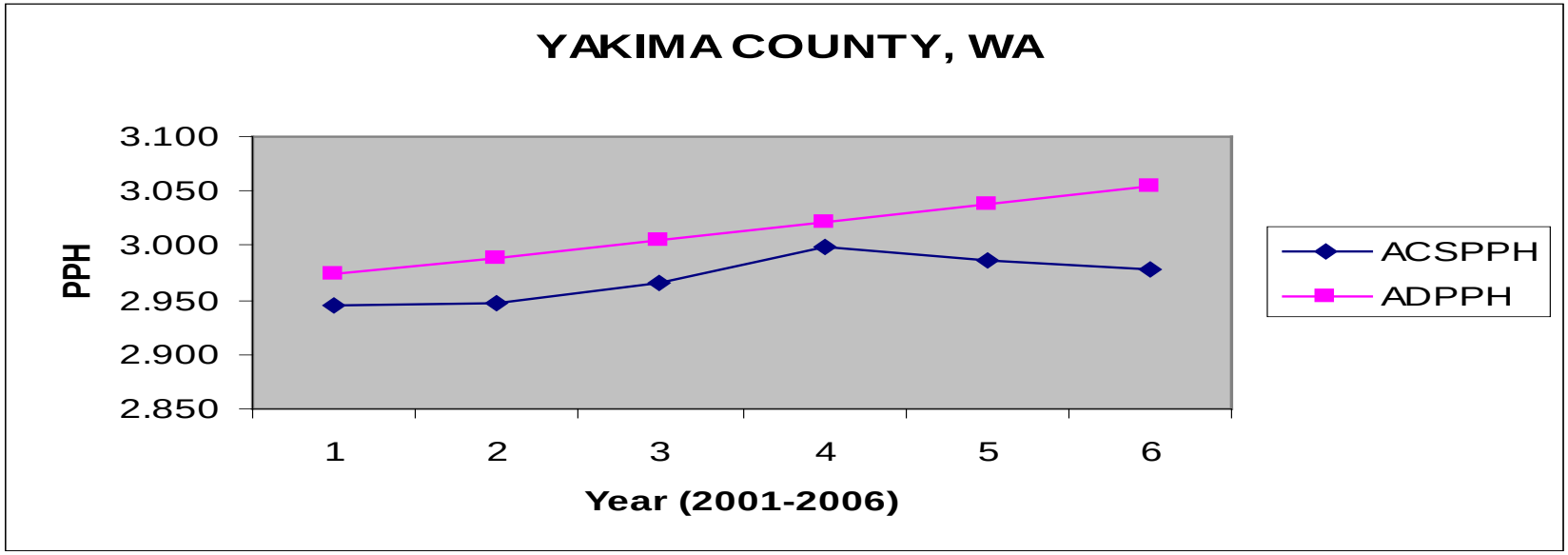

SCHUYKILL COUNTY, PA
PPH
2.380
2.360
2.340
2.320
2.300
2.280
2.260
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
AdPPH3YR

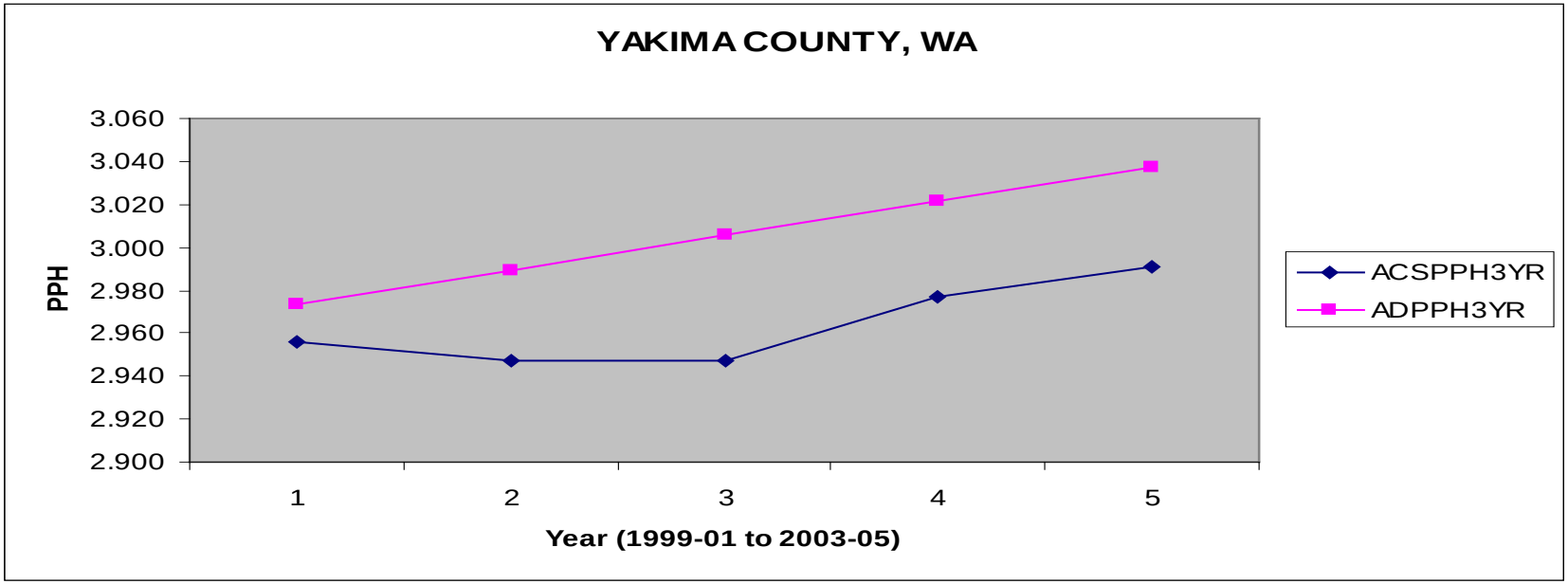
RESULTS


SEVIER COUNTY, TN
PPH
2.650
2.600
2.550
2.500
2.450
2.400
2.350
2.300
2.250
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH


SEVIER COUNTY, TN
PPH
2.480
2.460
2.440
2.420
2.400
2.380
2.360
2.340
2.320
2.300
2.280
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

RESULTS


YAKIMA COUNTY, WA
3.100
3.050
3.000
2.950
2.900
2.850
PPH
1
2
3
4
5
6
Year (2001-2006)
ACSPPH
ADPPH


YAKIMA COUNTY, WA
3.060
3.040
3.020
3.000
2.980
2.960
2.940
2.920
2.900
PPH
1
2
3
4
5
Year (1999-01 to 2003-05)
ACSPPH3YR
ADPPH3YR

RESULTS

In comparing the single-year ACS PPH values to the model-based PPH values, we find that the ACS PPH values are above the model-based PPH values in seven counties for the entire period, 2001-2006, that they are below the model-based values in two counties for the entire period and cross over the model-based values in nine counties (three of which have two crossovers each).

RESULTS

In terms of directional changes, the single-year ACS PPH values change direction three or more times in three counties, twice in nine counties, and once in five counties. Only one county (Blackhawk, Iowa) has no directional changes in its single-year ACS PPH values.

RESULTS

The three-year ACS PPH values remain above the model-based values for the entire period , 1999-2001 to 2003-2005 in nine counties, while in three counties they remain below the model-based values, and cross over the model based values seven times.

RESULTS

The three-year ACS PPH values change direction twice in two counties and once in seven counties. In the remaining nine counties no directional changes are observed. The model-based PPH values show a secular decline in 11 counties and an increase in seven.

RESULTS

In some of the counties with declining model-based PPH values, the trends are very slight (e.g., Pima County, Arizona) and in others, more distinct (Schuylkill County, Pennsylvania).

Similarly, some of the counties with increasing model-based PPH values have a very slight upward trend (e.g., San Francisco County, California), in others they are much more pronounced (e.g., Tulare County, California).

RESULTS

Table 2 provides mean PPH values across the 18 counties (and their standard deviations) by year. Not surprisingly, the single-year ACS PPH values exhibit the least systematic change over time and the most variation each year. In two of the six years, these values are lower than the model-based PPH values while in the remaining four years they exceed the model-based values. The means of the three-year ACS PPH values show a systematic decline over time with annual variations comparable to the model-based PPH values.

RESULTS

UCR University of California, Riverside

Table 2. Mean ACS Values by Year and Their Standard Deviations

Year	Mean 1-Year ACS PPH Values*	Mean Model-Based PPH Values*	Mean 3-Year ACS PPH Values*	Year
2001	2.503 (0.295)	2.627 (0.281)	2.648 (0.290)	1999-2001
2002	2.509 (0.287)	2.625 (0.286)	2.647 (0.286)	2000-2002
2003	2.642 (0.294)	2.622 (0.289)	2.642 (0.289)	2001-2003
2004	2.647 (0.319)	2.620 (0.300)	2.644 (0.300)	2002-2004
2005	2.623 (0.323)	2.618 (0.303)	2.635 (0.312)	2003-2005
2006	2.717 (0.312)	2.625 (0.309)	N/A	N/A

*The value shown in parentheses is the standard deviation (N=18)

DISCUSSION

As noted earlier, the US Census Bureau found encouraging results for the three-year ACS PPH values among the set of 1999 test counties when it compared the 1999-2001 numbers to the PPH values of the 2000 Census (US Census Bureau 2004b). As also was noted earlier, this finding was no surprise because the total household population and the total number of housing units found in Census 2000 are used as control variables in ACS weighting.

DISCUSSION

Given this, the results found here are a bit discouraging, given that these same variables are also used as control variables in ACS weighting – with one major change – once beyond the 2000 census, the total household populations and housing units are not enumerated directly, but, instead, estimated.

DISCUSSION

Not surprisingly, it is the single-year ACS PPH estimates that are the most discouraging. They jump around a great deal from year to year in many of the counties, a characteristic that is not desirable for both demographers who use the HUM and the stakeholders for whom HUM estimates are done.

DISCUSSION

This is because there is an expectation on the part of both these demographers and the stakeholders that PPH values should exhibit systematic changes unless there is compelling substantive evidence (e.g., the PPH values jumped because of a surge of in-migrants with high fertility and large family sizes) to the contrary.

DISCUSSION

If such PPH values are used in the absence of compelling substantive evidence justifying their temporal instability then it appears to me that the risk of challenges and related administrative and legal actions increases, especially when these estimates are used to allocate resources, which is often the case.

DISCUSSION

In considering the three-year ACS PPH values, the results are not as discouraging, as those for the single-year values, but neither are they strongly encouraging. These values change more systematically than do the single-year ACS PPH values, but they still exhibit temporal instability.

DISCUSSION

In addition to the temporal instability issue, one must ask what causes some of the substantial differences observed between the mean ACS PPH values and the mean model-based PPH values. For example, in 2001, the mean ACS PPH is 2.503 while the model-based mean is 2.627. This is a substantial difference, one likely beyond the scope of simple sampling error. Is this difference partly due to the ACS residency rule? It is not the same as the Decennial Census residency rule, the one that is inherent in the model-based ACS PPH values.

DISCUSSION

With a two-month rule, the ACS clearly will tend to have higher PPH values in areas in which seasonal migrants are currently residing than would be the case with the "majority of your time" rule used by the Decennial Census. This might explain in part the higher ACS PPH values found in Pima County, Arizona. However, if this were the case, we would expect that the ACS PPH values would consistently be higher than the model-based PPH values in Tulare County, California, but they are not.

CONCLUSIONS

- The single Year ACS PPH values viewed here do not appear to be useful to applied demographers who use the Housing Unit Method because of their fluctuation over time.
- The 3-Year ACS PPH values have less fluctuation than the single year values, but they do not appear to be much more useful than the single year values.

CONCLUSIONS

Our finding that the ACS PPH values are not particularly usable for purposes of making HUM-based population estimates is preliminary in nature.

More work needs to be done not only to confirm this finding, but given it is confirmed, we also need to figure out if the ACS PPH values can be modified so that they can be used.

FUTURE RESEARCH

With this in mind, our suggestions for further analysis include:

(1) conducting a broader scale comparison, taking into account the full range of counties;

(2) examining ACS PPH values that are not controlled;

(3) consideration of a way to utilize sample error (i.e., confidence intervals) in determining ACS PPH changes over time;

(4) an examination of 5-year ACS PPH values when at least three years of data become available; and

(5) making adjustments to ACS PPH values (deriving model-based PPH values from the ACS) that may provide more temporal stability.