UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

July 19, 2011

2011 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT MEMORANDUM SERIES #ACS11-RER-02

MEMORANDUM FOR	ACS Research and Evaluation Steering Committee
From:	Tommy Wright **{signed 7/19/2011}** Chief, Center for Statistical Research and Methodology
Prepared by:	Tucker McElroy Time Series Group Center for Statistical Research and Methodology
Subject:	Quantifying Discrepancies Between Diverse Multi-Year Estimates

Attached is the final American Community Survey Research and Evaluation report for "Quantifying Discrepancies Between Diverse Multi-Year Estimates." This project was conducted to answer questions raised by Alfredo Navarro of DSSD regarding multi-year estimates.

If you have any questions about this report, please contact Tucker McElroy (301-763-3227).

Attachment

cc:
ACS Research and Evaluation Team

American Community Survey Research and Evaluation Program

July 19, 2011

Quantifying Discrepancies Between Diverse Multi-Year Estimates

FINAL REPORT

Tucker McElroy, Natalya Titova, Chaitra Nagaraja
Center for Statistical Research and Methodology

Quantifying Discrepancies Between Diverse Multi-Year Estimates

Tucker McElroy[*], Natalya Titova[†], and Chaitra Nagaraja[‡]
U.S. Census Bureau

Abstract

The rolling sample methodology of the American Community Survey (ACS) introduces temporal distortions, resulting in Multiyear Estimates (MYEs) that measure aggregate activity over three or five years. They cannot be viewed as estimates of the final year, nor of the middle time point of the period; neither can they be viewed as simple averages of single year estimates belonging to the same time span. The U.S. Census Bureau has enunciated this principle forcefully, yet the question remains – in light of the fact that users are likely to ignore official pronouncements – how damaging are these unsanctioned viewpoints? In particular, can one quantify the impact of making such fallacious use of the MYEs? This paper answers these questions positively: yes, it can be quantified, and in general there is fairly serious degradation to the usability of MYEs when applying these faulty interpretations. We first offer a critique of the comparison of diverse MYEs via the published standard errors, and then we discuss a simple, general method based on relative percent discrepancies. This technique is illustrated on the test database of the ACS, from which we draw our general conclusions.

Keywords. American Community Survey, Rolling Sample, Sampling Error, Usability.

Disclaimer This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

1 Introduction

The American Community Survey (ACS) of the U.S. Census Bureau was designed as a more timely analogue of the Census Long Form, with data on social, demographic, economic, and housing

[*]Statistical Research Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

[†]Statistical Research Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, natalya.titova@census.gov

[‡]Statistical Research Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, chaitra.nagaraja@census.gov

variables being published continuously every year for a range of geographical regions. In order to counteract perceived high sampling variability in geographical regions of low population, it was deemed prudent to aggregate information temporally using a rolling sample. One consequence of this approach is a decreased utility of the data for the purposes of temporal analysis, since the rolling sample has the effect of retarding the timeliness of trends and other temporal characteristics. For a background and discussion of these issues, see the following: Kish (1981), Alexander (1998), Citro and Kalton (2007), Beaghen and Weidman (2008), McElroy (2009), Beaghen, Weidman, Asiala, McElroy, and Navarro (2010).

The rolling sample methodology was selected for its simplicity and its reduction of sampling variability (Citro and Kalton (2007)). However, as it was recognized that this method would generate lag in underlying trends and temporal dynamics (cf. Bell (1998) and Breidt (2007)), the concept of a period-estimate was brought forth. Thus, a rolling sample estimate over a span of years was not to be viewed as an estimate of the end year or the middle year, but simply as an estimate of aggregate activity over the entire temporal span. This approach is mathematically viable and entirely analogous to the concept of flow estimates used in economic time series.

The resulting period-estimates, formally called Multiyear Estimates (MYEs), are currently being published for counties, tracts, block groups, and school districts throughout the U.S.A (including territories), with periods of one-, three-, and five-years. The temptation is to view a particular MYE published in a given year as an estimate of activity *in that given year*. This is a fallacy. If one were to measure the temperature on July 1 by averaging temperature readings over the entire year, one would not expect it to be a reliable estimate unless temperature in that region was stable over time. More properly speaking, one has an estimate that is proportional to aggregate temperature over the entire year. Likewise, an MYE represents aggregate activity over several months; the U.S. Census Bureau has gone to great lengths to communicate this correct interpretation.

Part of the intellectual difficulty is that a three-year or five-year period is not an interval time associated with any culturally significant epoch in the human mind. For example, a 10-year MYE could be described as the "decade estimate," the term "decade" at once conjuring the important consideration that the figure represents activity over the entire period of time, not just the final year. No such temporal mnemonic exists for three- and five-year periods.

Given that the average consumer of the ACS data may be inclined to disregard official U.S. Census Bureau pronouncements on the proper use and interpretation of its products, and may indeed fallaciously proceed to view MYEs as estimates of the middle or the end year, what are the consequences? Can these consequences be quantified in a statistical fashion? We seek in this paper to provide a methodological framework to answer this question, as well as to provide extensive empirical evidence from the ACS that quantifies these consequences. A related question arises from the consideration that an MYE represents aggregate activity over a period of time; can MYEs be viewed as simple averages of annual estimates (or 1-year MYEs) drawn from the corresponding years? Intriguingly, this type of linear relationship is implicitly assumed in the material of Citro and Kalton (2007) and U.S. Census Bureau (2007), where the authors supply instructions on the computation of sampling error variances for linear combinations of MYEs. A negative answer to our query therefore may shed doubt on the usefulness of this approximation.

Our paper first introduces some notation in Section 2 along with a brief discussion of the ACS data. The frame of our empirical analysis is the Multiyear Estimates Study (MYES), a pre-flight version of the ACS restricted to a set of counties in the USA. Section 3 discusses the main questions of our study in the context of the statistical behavior of sampling errors. We explain why sampling errors across MYEs of different period-lengths tend to be correlated, and this serves as the motivation for our method in Section 4. This method of evaluation is probabilistic rather than statistical in nature, and we justify it on the basis of usability and interpretability of the MYEs. Actual empirical results are found in Section 5, applying our methodology to the MYES. The vast amount of information is summarized in several tables, and specific relationships of interest are highlighted through an in-depth discussion of selected variables. Section 6 summarizes our conclusions.

2 The ACS Data

We will denote an MYE by $X_t^{(k)}$ for $k = 1, 3, 5$, where t is the final year of the period (which is typically not the same as the year of publication) and k refers to the number of years entering the rolling sample. We use the term "period" to denote a stretch of time, e.g., the 2003–2005 period estimate refers to the 3-year MYE covering years 2003, 2004, and 2005. Clearly the $k = 1$ case is not strictly speaking a multiyear estimate, but is nevertheless a period estimate; however, for ease of exposition we will maintain this abuse of official terminology.

As mentioned in Section 1, when the population of a given region is sufficiently low, one-year estimates may not be produced; data may have to be aggregated across years. MYEs are needed when the population of a given region is sufficiently low, i.e., less than 65,000 people. If the population is 20,000 or more three years of data are utilized, whereas for all regions five years of data are pooled. For instance, state-level estimates are published separately for 1-, 3-, and 5-years worth of data, whereas a census tract may only have five-year estimates published if the population is below 20,000. See U.S. Census Bureau (2006) and Torrieri (2007) for more discussion. If we use data from the Multiyear Estimates Study (MYES), we can obtain contiguous time series data going back to 1999, which is not the case for current ACS estimates. In practice, the data from the MYES and the current ACS are extremely similar; there are only few minor differences in methodology[1] (e.g., group quarters were not included in the MYES).

The MYES was a trial study for the ACS restricted to 34 counties, utilizing data from 1999

[1] Navarro (2009).

through 2005. However, the actual MYEs that were published are: 2000 through 2005 for 1-year; 2001 through 2005 for 3-years; 2003 through 2005 for 5-years. Since we are interested in making comparisons between MYEs of different period length, it makes sense to consider only those pairings drawn from the same underlying spans of time. This principle produces four comparisons of 1-year to 3-year MYEs: $\{X^{(1)}_{2000}, X^{(1)}_{2001}, X^{(1)}_{2002}\}$ with $X^{(3)}_{2002}$; $\{X^{(1)}_{2001}, X^{(1)}_{2002}, X^{(1)}_{2003}\}$ with $X^{(3)}_{2003}$; $\{X^{(1)}_{2002}, X^{(1)}_{2003}, X^{(1)}_{2004}\}$ with $X^{(3)}_{2004}$; and $\{X^{(1)}_{2003}, X^{(1)}_{2004}, X^{(1)}_{2005}\}$ with $X^{(3)}_{2005}$. There are also two possible comparisons of 1-year to 5-year MYEs: $\{X^{(1)}_{2000}, X^{(1)}_{2001}, X^{(1)}_{2002}, X^{(1)}_{2003}, X^{(1)}_{2004}\}$ with $X^{(5)}_{2004}$; and $\{X^{(1)}_{2001}, X^{(1)}_{2002}, X^{(1)}_{2003}, X^{(1)}_{2004}, X^{(1)}_{2005}\}$ with $X^{(5)}_{2005}$. Other comparisons are possible, of course, and are considered in Section 4.

Potentially the time series, at least for these 34 counties, could be concatenated with current ACS estimates to form a longer sample, but unfortunately this is impossible due to the current publication schedule. That is, the current ACS estimates (for all of the U.S.A.) began publication with 1-year MYEs in 2006, 3-year MYEs in 2008, and 5-year MYEs in 2010. This leaves some gaps in the time series data, which in McElroy (2009) was resolved through the crude device of forecasting. In this paper we focus instead on the MYES, with the assumption that results for the trial period should also hold valid for the nation at large (note that the counties for the MYES were selected by experts in the belief that they constituted a diverse and representative picture of the whole U.S.A.).

For the actual construction of MYEs, see the discussion in Fay (2007), Starsinic and Tersine (2007), Tersine and Asiala (2007), and McElroy (2009). The procedure is complicated, involving sampling weights, nonresponse adjustment, and population controls; the result is a highly nonlinear operation on the samples, further interfering with the linear picture adopted in Citro and Kalton (2007). Nevertheless, it may be the case that these linearity-destroying operations are not practically significant, or at least do not have an impact on the interpretation and use of the MYEs.

For each of the MYEs in a given year, there are data available at the county aggregation level for each of 34 counties. However, the data for $X^{(1)}_{2000}$ are only available for 19 counties, so we restrict to these – see Table 3 for a listing with codes. There are hundreds of variables available, which are partitioned into Economic, Housing, Social, and Demographic categories. Some of these figures are Numerics (e.g. totals, rates, averages), whereas others are Percents (i.e., proportions); in each case sampling variances are published. An important note is needed here: the literature on this topic typically assumes that all statistical variation is due to sampling mechanisms. This approach is not tenable for economic and demographic time series data; it is more common in published literature to adopt the view that sampling error is but one component of a sampling estimate, with other sources of stochastic variation potentially present in the underlying population. Therefore, references in the literature to "variances" and such should have the adjective "sampling" inserted before them as an essential qualifier. Nevertheless, in our study we are principally interesting in sampling variation, and we will therefore ignore other stochastic dynamics in the data.

3 Methodology Utilizing Sampling Standard Errors

In this section we discuss issues connected with how MYEs might be compared utilizing the statistical properties of sampling errors. In Citro and Kalton (2007) these standard deviations of the sampling errors are used to compute confidence intervals for population values. As discussed in the Introduction, we are interested in quantifying the following questions:

1. Is the k-year MYE ($k = 3, 5$) different from the average of corresponding 1-year MYEs?

2. Is the k-year MYE ($k = 3, 5$) different from the 1-year MYE corresponding to the middle year?

3. Is the k-year MYE ($k = 3, 5$) different from the 1-year MYE corresponding to the final (or end) year?

We show here that a negative answer to the first question is implicitly assumed in the methodologies espoused in Citro and Kalton (2007)[2]. Similar issues are explored in Sections 5.5, 5.6, and 6.3 of Beaghen and Weidman (2008)[3]. If we conceive of inference as pertaining to true population quantities, then we may define the sampling errors to be the difference between the estimates and these true population quantities. This provides the following simple description:

$$
\begin{aligned}
X_t^{(1)} &= x_t^{(1)} + \epsilon_t^{(1)}, \\
X_t^{(3)} &= x_t^{(3)} + \epsilon_t^{(3)}, \\
X_t^{(5)} &= x_t^{(5)} + \epsilon_t^{(5)}.
\end{aligned}
$$

The quantities $x_t^{(k)}$ are the population values, and from the perspective of sampling theory can be viewed as deterministic. The $\epsilon_t^{(k)}$ are sampling error variables, and are stochastic; typically their randomness can be viewed as completely determined by the sampling mechanism. Without loss of generality, they have mean zero, and their variances are estimated via the sampling variances $V_t^{(k)} = \widehat{Var}[\epsilon_t^{(k)}]$. In U.S. Census Bureau (2007)[4] – from which Beaghen and Weidman (2008) draw their methodology – it is implicitly assumed that the sampling errors may be viewed as having a normal distribution with the exact variance given by $V_t^{(k)}$. Given the size of samples typically used, these assumptions seem reasonable, being founded on basic theoretical ideas of the sampling literature.

[2]This work discusses how to compute standard errors for overlapping MYEs, i.e., those MYEs of common period lengths that share some years in common.

[3]Tables 5 and 7 provide a "SE of Difference" between a 1-year and 5-year (overlapping) MYEs. The method used is not described, but presumably they follow the technique of U.S. Census Bureau (2007) that assumes a linear form for the MYEs.

[4]Page 8 of the official document gives confidence interval widths in terms of the standard errors, which exactly correspond to the Gaussian quantiles.

Although it would be convenient to suppose that the sampling errors are not serially correlated, and that they are not cross-correlated with regard to period length, this is an untenable assumption in general (though for period estimates drawn from non-overlapping years, it is reasonable to presume the sampling errors are either uncorrelated or have a small negative correlation). In particular, assuming that sampling errors for different period lengths – covering overlapping time periods – are independent is not realistic since they are all derived from the same pool of data. We illustrate this problem with a simple example.

Suppose the population value for a variable of interest in year t is x_t, and in that year we collect a sample of size m_t: $\{X_{t,1}, X_{t,2}, \ldots, X_{t,m_t}\}$. Then, $X_{t,i} = x_t + e_{t,i}$ for sampling errors $\{e_{t,i}\}$ which are independent and identically distributed according to the sampling mechanism. Let us consider years $t = 1, 2, 3$. With a simple, equally-weighted sampling scheme, an average effect, such as mean travel time to work, would be estimated for each year t using

$$X_t^{(1)} = \frac{1}{m_t} \sum_{i=1}^{m_t} X_{t,i} = x_t + \frac{1}{m_t} \sum_{i=1}^{m_t} e_{t,i} = x_t + e_t^{(1)}.$$

The 3-year MYE for $t = 3$ would then be estimated by

$$X_3^{(3)} = \frac{1}{m_1 + m_2 + m_3} \sum_{t=1}^{3} \sum_{i=1}^{m_t} X_{t,i}. \tag{1}$$

Now, further suppose that all sample sizes are equal. That is, $m_1 = m_2 = m_3 = m$. Then, we can rewrite (1) as follows:

$$X_3^{(3)} = \frac{1}{3} \left(\frac{1}{m} \sum_{i=1}^{m} \sum_{t=1}^{3} X_{t,i} \right) = \frac{1}{3} \sum_{t=1}^{3} x_t + \frac{1}{3} \sum_{t=1}^{3} e_t^{(1)}. \tag{2}$$

Thus by definition of $X_3^{(3)}$ we have $x_3^{(3)} = \frac{1}{3} \sum_{t=1}^{3} x_t$ and $e_3^{(3)} = \frac{1}{3} \sum_{t=1}^{3} e_t^{(1)}$, and in particular $e_3^{(3)}$ and $e_3^{(1)}$ are cross-correlated. This type of calculation is at the heart of the method suggested by the U.S. Census Bureau (2007, page 12). However, (2) cannot be correct in general for the following reason. It implies that $\frac{1}{3} \sum_{t=1}^{3} X_t^{(1)} = X_3^{(3)}$, or in other words that our first question raised at the start of this section is always true. But in practice, actual computation of the difference of these quantities yields non-zero values, which is documented below. Many factors contribute to the falsity of the above description of MYEs: variations in sample sizes over year, weighting patterns, non-response adjustment, controls, etc. To summarize, we have made the following observations:

- Sampling errors for an MYE viewed over time are serially correlated, being temporally k-dependent ($k = 1, 3, 5$).

- Sampling errors for MYEs of different period-lengths are cross-correlated if the time periods overlap.

- A strictly linear representation of MYEs – as in U.S. Census Bureau (2007) and Beaghen and Weidman (2008) – is not useful for answering our questions, since it presumes a relationship that is typically false.

With regard to this last point, we note that it may be the case that the actual differences are so small that taking a linear representation presents no real loss of utility. That is, even though the linear representation is false as an exact equality, it may be a suitable approximation in the sense that there is no negative consequence associated with its use. So we might modify the three questions supplied at the start of this section, by inserting the adverb "consequentially" before the word "different." (We avoid the adverb "significantly," since this conjures concepts of statistical significance; we wish to consider broader implications with the word "consequentially.") Thus we see that the ideas and assumptions that underpin U.S. Census Bureau (2007) and Citro and Kalton (2007) cannot be used as the basis of our methodology; in order to establish the general validity of their framework, another technique is needed. We describe it in the next section.

4 Methodology Utilizing Relative Percent Differences

The discussion of Section 3 provides the motivation for a method of testing for "consequential differences." In McElroy (2009) discrepancies in MYEs were assessed through relative percent differences defined via

$$D_{ave,t}^{(k)} = \frac{\left(\sum_{j=0}^{k-1} X_{t-j}^{(1)}\right)/k - X_t^{(k)}}{X_t^{(k)}} \quad \text{for } k = 3, 5;$$

$$D_{mid,t}^{(k)} = \frac{X_{t-(k-1)/2}^{(1)} - X_t^{(k)}}{X_t^{(k)}} \quad \text{for } k = 3, 5;$$

$$D_{end,t}^{(k)} = \frac{X_t^{(1)} - X_t^{(k)}}{X_t^{(k)}} \quad \text{for } k = 3, 5.$$

More generally, we have $D_{lin,t}^{(k)} = \overline{X}_{lin,t}^{(k)}/X_t^{(k)} - 1$, where *lin* denotes a linear estimate of $X_t^{(k)}$ based on one-year estimates, being either *ave*, *mid*, or *end* depending on the case under consideration. These quantities can be computed for all variables, counties, and available times. However, we do not have any distributional information about these quantities – since it is unlikely that they are independent and identically distributed across region, variable type, and time[5], the empirical distribution is not meaningful. These $D_{lin,t}^{(k)}$ simply measure the distortion arising from substituting an MYE $X_t^{(k)}$ with the corresponding $\overline{X}_{lin,t}^{(k)}$. If one makes this substitution, what is the effect on our

[5]Regional patterns are likely to be mirrored in local and global patterns for certain variables; some variables are strongly linked, as they have aggregation relations and dependencies built into their definitions. Temporal correlation is to be expected due to the trending nature of many economic and demographic variables.

confidence about the true population value? The ACS publishes standard errors that are advocated (U.S. Census Bureau (2007)) as being useful for the following types of inference[6]:

$$0.95 \approx \mathbb{P}\left[x_t^{(k)} \in \left\{X_t^{(k)} \pm 1.96\sqrt{V_t^{(k)}}\right\}\right].$$

That is, the confidence interval covers the true population value, determined by the sampling estimates and their standard errors, with an approximate probability of 95%. In this paper, we are only interested in confidence intervals for $k = 3, 5$. This interval can be re-written as

$$\begin{aligned}&\left[\overline{X}_{lin,t}^{(k)} - D_{lin,t}^{(k)} X_t^{(k)} - 1.96\,\sqrt{V_t^{(k)}},\ \overline{X}_{lin,t}^{(k)} - D_{lin,t}^{(k)} X_t^{(k)} + 1.96\,\sqrt{V_t^{(k)}}\right] \\ &= \left[\overline{X}_{lin,t}^{(k)} - X_t^{(k)}\left(D_{lin,t}^{(k)} + 1.96\,CV_t^{(k)}\right),\ \overline{X}_{lin,t}^{(k)} - X_t^{(k)}\left(D_{lin,t}^{(k)} - 1.96\,CV_t^{(k)}\right)\right].\end{aligned}$$

Here $CV_t^{(k)}$ stands for the coefficient of variation $\sqrt{V_t^{(k)}}/X_t^{(k)}$. If $D_{lin,t}^{(k)} = 0$, then we can swap $\overline{X}_{lin,t}^{(k)}$ for $X_t^{(k)}$ with no loss, since the confidence interval is unchanged. Likewise, if $|D_{lin,t}^{(k)}|$ is quite small relative to $CV_t^{(k)}$, the confidence interval will be changed very slightly. Writing $|D_t^{(k)}| = \delta CV_t^{(k)}$ for some small δ, the question is: how large can δ be such that the resulting interval is still approximately 95%? The approximate coverage – utilizing the normal distribution – is that of a skewed interval $\overline{X}_{lin,t}^{(k)} + (-\delta \pm 1.96)\sqrt{V_t^{(k)}}$, or

$$\Phi(-\delta + 1.96) - \Phi(-\delta - 1.96) = g(\delta).$$

The function $g(\delta)$ gives the real asymptotic coverage when making the substitution; note that $g(0) = 0.95$, the maximum value[7]. We propose to choose δ such that the resulting coverage is different from 95 percent by less than one half a percentage point (call this value τ). This threshold is chosen so that if one rounded the coverage to the nearest integer percentage, there would be no difference. This means finding the maximum $|\delta|$ such that $g(\delta) \geq 0.945$. The value $|\delta| = 0.208$ satisfies this (to three decimal places), in the sense that $\delta \geq 0.209$ will yield $g(\delta) < 0.945$.

This principle can of course be generalized beyond the $\alpha = 0.05$ significance level and the $\tau = 0.005$ maximum discrepancy amount. Table 1 produces the relation between coverage level α and the maximal $|\delta|$, while Figure 1 illustrates the latter's dependence on τ. They serve to demonstrate the flexibility of our method. However, our empirical analysis in the sequel focuses on the $\alpha = 0.05$, $\tau = 0.005$ case. In summary, when $|D_{lin,t}^{(k)}|$ exceeds $0.208\,CV_t^{(k)}$ we deem there to be a "consequential" discrepancy between $X_t^{(k)}$ and $\overline{X}_{lin,t}^{(k)}$, in the sense that the confidence interval coverage would be altered nontrivially if one made the substitution. Therefore we adopt the following rule:

[6] The above reference uses the value 2 instead of 1.96 for the 95 % confidence interval.

[7] Proof: $\dot{g}(\delta) = \phi(-\delta - 1.96) - \phi(-\delta + 1.96)$ so that $\dot{g}(0) = 0$ by the symmetry of ϕ; hence $\delta = 0$ is a critical point. But $\ddot{g}(0) = \dot{\phi}(1.96) - \dot{\phi}(-1.96) = -2(1.96)\ \exp -(1.96)^2/2/\sqrt{2\pi} < 0$, which implies a local maximum.

$1-\alpha$	δ
0.900	0.171
0.950	0.208
0.966	0.235
0.975	0.261
0.980	0.281
0.983	0.298
0.985	0.311
0.987	0.327
0.988	0.336
0.990	0.359

Table 1: For each given α we report the value of δ that is maximal such that the resulting discrepancy in the confidence interval is at most $\tau = 0.005$.

$$D_{lin,t}^{(k)} \text{ is consequential if } \left| D_{lin,t}^{(k)} \right| > 0.208 \times \sqrt{V_t^{(k)}} / X_t^{(k)}. \tag{3}$$

5 Results

The previous section sets forth our methodology. This can easily be applied to any variable, in any region, at any time – so long as both types of MYEs under comparison are available. We now draw some general conclusions about comparability of MYEs via this method, based upon the MYES database discussed in Section 2. We provide summaries of our results from this large database – due to the large number of variables, it is infeasible to provide individual results. In Sec. 5.1 we identify which relative percent differences are consequential, using the δ-criteria described previously. Then Sec. 5.2 examines twenty variables in greater detail; these variables were chosen to represent the wide range of types of variables found in the data, including totals for subpopulations, means, and income. In Sec. 5.3, we look at the results in finer detail for a specific variable and county combination instead of in aggregate as in Sec. 5.1. Finally, in Sec. 5.4, we examine whether the results change when we vary the maximum permitted discrepancy τ.

5.1 Summary of Results

Recall that the notion of a consequential relative percent difference is defined in Section 4 via comparison with $\delta = 0.208$. The comparisons are summarized in the following manner: first fix a time period (i.e., a span of either 3 or 5 years, depending on the MYE type considered), and

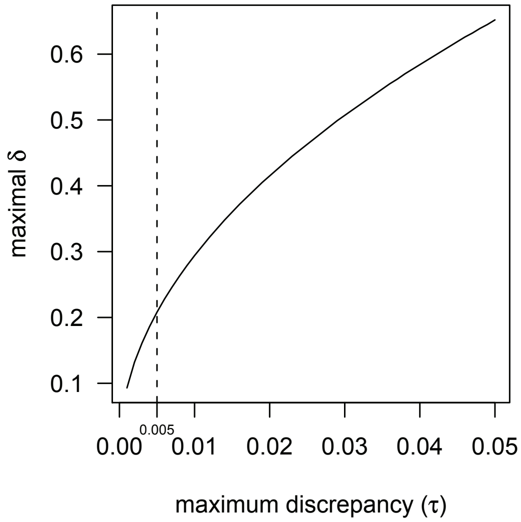

Figure 1: This figure plots the allowed discrepancy τ to the confidence interval versus the resulting maximal δ, for a two-sided .95 interval.

consider all variables within a given county. These variables can be partitioned among Numerics and Percents (defined in Section 2), as well as category (i.e., Demographic, Social, Economic, or Housing). We consider relative percent differences for a variable type (Numeric or Percent) as well as a variable category for each county; if for a particular county we have a total of y available relative percent differences, out of which x of them are consequential, then the proportion of consequential relative percent differences is x/y. There are 19 such ratios (one for each county), and we further summarize them by reporting both the average and the maximum value. For example, according to Table 4, the average of 19 proportions (from 19 counties) of consequential relative percent differences for the *mid* comparison (i.e., $|D^{(3)}_{mid,2002}| = |X^{(1)}_{2001} - X^{(3)}_{2002}|/X^{(3)}_{2002}$) – where the $X^{(k)}_t$s are numerics – is 0.91. Also, the maximum of the same 19 proportions of consequential relative percent differences is 0.97.

The approach of Section 4 presents a very low threshold – relative to the values that are typically observed – in order for a relative percent difference to be deemed consequential. That is why the proportions in Tables 4 and 5 are very large, i.e., most relative percent differences are consequential. Of course, setting the amount of distortion to the confidence interval to at most 0.5% requires a very small difference; tolerating a greater distortion – or changing the α level – would alter these results. Recall that Table 1 provides alternative δ values for different α. For example, with 99% coverage a confidence interval is much wider, so that more distortion can be ignored, essentially allowing for the higher δ value of 0.359; adopting this δ-rule instead would reduce the rate of consequential discrepancies, since they tend to incur less damage to usability.

According to Table 4, among Demographic and Housing variables, the $|D^{(3)}_{ave,t}|$s have the lowest average and maximum proportions for Numerics, and the $|D^{(3)}_{end,t}|$s have the highest average and maximum proportions for Numerics, but the results are rather mixed for Social and Economic variables. Among all four types of variables, the $|D^{(3)}_{ave,t}|$s always have the lowest average and maximum proportions for Percents; the $|D^{(3)}_{end,t}|$s have the highest average and maximum proportions for Percents, with few exceptions. Proportions tend to be smaller for Percents than for Numerics, but this is not always true.

According to Table 5, among all four types of variables, the $|D^{(5)}_{ave,t}|$s always have the lowest average proportions for Numerics, and the $|D^{(5)}_{end,t}|$s always have the highest average proportions for Numerics, but maximums are very similar and slightly lower for $|D^{(5)}_{mid,t}|$s. The $|D^{(5)}_{end,t}|$s always have the highest average proportions for Percents, but maximums are very similar and slightly lower for $|D^{(5)}_{mid,t}|$s. Average proportions for Percents are the lowest for $|D^{(5)}_{ave,t}|$s among Economic and Housing variables, and for $|D^{(5)}_{mid,t}|$s among Demographic variables.

5.2 Examining Specific Variables

Here we examine more closely 20 specific variables of interest. These are listed in Table 6 along with the table and profile line numbers for each. The variables were chosen to represent a variety of "types" found in the ACS data, which we describe below. Some of the variables are included in multiple categories[8].

1. **Stable variables**: There are few variables where we would not expect large differences in a three- or five-year period. For such variables, we would expect some, if not all, of the three interpretations of the period estimate to be reasonable. We examine three such variables here: number of males (DP01/2), average household size (DP01/103), and fertility per 1,000 unmarried women (DP02/31). Additional variables listed below in the means/medians/rates category may also fit this description.

2. **Subpopulations**: This is a key situation wherein estimates can differ substantially from year to year, because very few people may be sampled who exhibit a given characteristic[9]. Therefore, interpreting a MYE to be an estimate of the middle or end of the period may not be appropriate for such variables. The list of variables selected specifically for this characteristic: number of Filipino residents (DP01/42); number of grandparents who are responsible for their grandchildren (DP02/37); veteran status/civilian veterans (DP02/41); number of unpaid family workers (DP03/49); number of rooms in home is 5 (DP04/26); number of people whose house heating fuel is solar energy (DP04/59).

[8]We refer to each variable as follows: (table number/profile line). For instance, median age is represented by DP01/17.

[9]A specific example in a related context can be found in Beaghen and Stern (2009).

3. **Means/Medians/Rates**: Most of the Numeric variables are totals; far fewer are means, medians, or rates. Therefore, it is important to determine whether conclusions for such variables differ from other types of Numeric variables. We expect these values to be fairly stable. The variables in this category are: median age (DP01/17); average household size (DP01/103); fertility per 1,000 unmarried women (DP02/31); mean travel time to work (DP03/25); mean household income (DP03/62); homeowner vacancy rate (DP04/4).

4. **Money-related variables**: Questions regarding money are in a category of their own, not only because it is a sensitive question (see below) but also since the survey answers are adjusted for inflation and are used in computing poverty estimates (Beaghen, et al, 2010). The variables are income $10,000 and under (DP03/51); income between $35,000-$49,000 (DP03/55); mean household income (DP03/62); value of housing unit between $150,000-$199,000 (DP04/72); selected monthly house owner costs as a percentage of household income is 35% or higher (DP04/99); percent of all families in poverty.

5. **Sensitive questions**: There are some questions in the ACS survey that respondents may answer with greater hesitancy. They may not even know an answer to the question. It may be the case that these questions are left blank, or are answered with the respondent's best guess. Therefore, it is possible that these estimates differ substantially across time. The list of variables is: not a U.S. citizen (DP02/59); speaks English less than "very well" (DP02/79); income $10,000 and under; income between $35,000-$49,000; value of housing unit between $150,000-$199,000; selected monthly house owner costs as a percentage of household income is 35% or more.

In Table 7, proportions of relative percent differences that were found to be consequential are listed for each of the 20 selected variables. In the left part of Table 7 the maximum number of percent differences available is 76 (19 counties times 4 three-year studies). In the right part of Table 7 the maximum number of differences available is 38 (19 counties times 2 five-year studies).

Overall, the results from each variable seem to be similar. Most often, the end interpretation of the period estimate is the least acceptable, followed by the middle and average interpretations. There are exceptions as well: for example, for mean household income (DP03/62) the end interpretation of the period estimate is the most acceptable compared with other interpretations – perhaps because income is inflation adjusted.

The proportion of consequential relative percent differences seems generally to be slightly greater for each of the specific variables when compared with the results with all of the variables from the respective variable category (i.e., Demographic, Social, Economic, or Housing). A notable exception is the results for the average interpretation for the three-year estimates. Especially for the variables that were classified as subpopulations, the proportion of consequential relative percent differences is much lower than for the the overall results. It is also interesting to note that variables categorized as stable, such as fertility rates, violate the linearity assumptions at nearly the same rates as other variables.

If the three- and five-year results are compared with each other, we find that the fraction of consequential relative percent differences is higher for five-year estimates than for three-year estimates. This matches the results obtained using the full set of variables. While we have no rigorous explanation for this phenomenon, we speculate that the greater length of the five-year MYE provides more opportunities for linearity-destroying operations to take effect.

5.3 Comparing interpretations for Franklin County, OH

In this section we examine the results for an individual case: the number of males living in Franklin County, OH. We will show both numerically and graphically how the procedure introduced in Sec. 3 is implemented and interpreted. The 3-year estimates will be compared with the average, middle, and end interpretations constructed using the corresponding 1-year estimates.

In Table 2, the 3-year estimates (and variances) are provided along with the average, middle, and end interpretation percent differences $D^{(3)}_{lin,t}$ for each period[10]. The threshold value, as defined in (3), is provided for each period as well. For each interpretation and period combination, $|D^{(3)}_{lin,t}|$ is greater than the threshold $\delta CV^{(3)}_t$ and therefore all differences are consequential. Consequently, none of the interpretations is a suitable way to describe a 3-year estimate in this specific case.

We can see this even more clearly in Fig. 2. Here, we plot the 3-year estimates in red and the three interpretations in black[11]. A 95% pointwise confidence band is plotted around the 3-year estimates to show the variance around each estimate. In addition, the 1-year estimates along with the pointwise 95% confidence band is also plotted. Visually, we can see a large discrepancy between the red and any of the black lines and even the 1-year estimates themselves. Even though the differences between the lines amount to only a few thousand, which is a small percentage of the estimated population size, the trends lines for each interpretation differ not only from each other but also from the 3-year estimates. For instance, the middle interpretation of the 3-year estimate is trending down at the end while the others are moving up. Therefore, as in Sec. 5.1, we have evidence to show that none of these interpretations is appropriate.

5.4 Varying τ

Up until this point, we have fixed the maximum discrepancy (τ) to be 0.005, where $\alpha = 0.05$. One could argue that this threshold is too strict and that it is unsurprising most relative percent differences are regarded as consequential. If we increase the maximum discrepancy τ, the fraction

[10] The variances were truncated to be whole numbers; the decimal values were also truncated to have only 3 significant digits.

[11] Note that the listed year on the x-axis represents the *final* year in the 3-year period. E.g., the 3-year estimates containing data from 2000-2002 would be be plotted at year 2002.

Table 2: Comparing interpretations: Number of males in Franklin County, OH

	Year			
	2000-2002	2001-2003	2002-2004	2003-2005
3-year estimate: $X_t^{(3)}$	515,173	518,560	519,571	521,467
3-year estimate variance: $V_t^{(3)}$	4,364	6,303	14,545	6,694
$\delta CV_t^{(3)}$	2.66e-05	3.18e-05	4.82e-05	3.26e-05
average: $D_{ave,t}^{(3)}$	-4.23e-03	-2.35e-03	2.88e-03	2.38e-03
middle: $D_{mid,t}^{(3)}$	-8.68e-03	1.05e-03	5.06e-03	8.39e-04
end: $D_{end,t}^{(3)}$	7.63e-03	7.02e-03	4.49e-03	4.91e-03

$\alpha = 0.05$, $\tau = 0.005$, and $\delta = 0.208$ for this example.

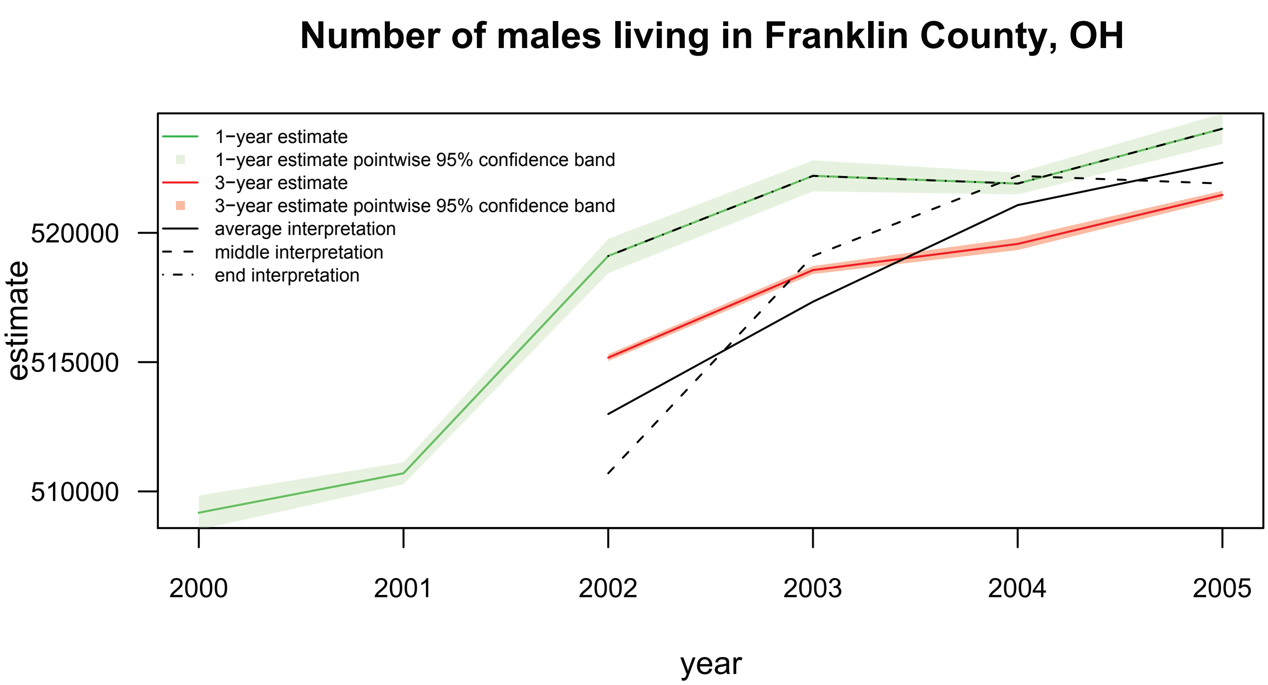

Figure 2: A visual comparison of interpretations for 3-year estimates

of consequential differences will decrease for any of the three interpretations. In this section, we examine how much our results actually change if we choose alternative values of τ (keeping α fixed at 0.05).

To answer this question, we perform an analysis similar to the one in Sec. 5.1 for four counties at varying levels of τ. The four counties, Franklin County, OH, San Francisco County, CA, Black Hawk County, IA, and Madison County, MS were chosen to represent the wide range of county population sizes (listed in descending order here). In Fig. 3, we plot the fraction of consequential differences for the set of numeric social variables (DP02) for each interpretation of the 5-year estimates for 2000-2004. As a benchmark, $\tau = 0.005$, which we used for the results in Sec. 5.1, is highlighted on each plot as well.

For each county, as the maximum permitted discrepancy (τ) increases, the proportion of consequential differences drops off much faster for the "average" interpretation than for the "middle" and "end" interpretations. For the latter two interpretations even at $\tau = 0.05$, more than 80% of the variables are consequentially different for any of the counties. Even at $\tau = 0.05$, except for Madison County, MS, more than 60% of the variables in each county are considered to be consequential. Even for large values of τ, no alternative interpretation of the 5-year estimate seems suitable. The same conclusion holds if we examine the corresponding plots for demographic (DP01), economic (DP03), and housing (DP04) variable sets.

6 Conclusion

This paper addresses a long-standing issue of concern in the ACS: can MYEs be viewed as linear functions of one-year estimates? Although official U.S. Census Bureau policy answers this question negatively, an empirical exploration of the issue has been lacking. We propose a simple replicable methodology that is directly matched to the typical user's concerns via the novel concept of "consequential discrepancy," and apply this technique to the MYES database. Our conclusion is that the official policy is correct. Note that fundamental reasons can be given for why MYEs should not be viewed as linear functions of one-year estimates, but even so it is not *a priori* clear that ignoring this injunction will have negative repercussions in terms of data analytical conclusions. However, the results given in Section 5 demonstrate that there are indeed negative consequences in practice.

Our methodology (Section 4) is necessarily simplistic. We cannot utilize methods involving statistical significance due to the noted cross-dependence of sampling error components corresponding to different period-length MYEs (see Section 3). We have also purposely avoided any in-depth attempt to model the MYES and mine it for patterns. One reason for this is the dearth of temporal data, i.e., we can only observe a MYE across a very limited time horizon when using data from the MYES. While interesting relationships may (and do) exist across variable type and county, these patterns cannot serve to address the question of cross-period comparability. The method of relative percent differences and the interpretation of consequential discrepancies is directly linked to underlying assumptions and usages of the ACS already in place, and adds no further axiomatic burden. There are few "tuning parameters" to the method, and the default values used in the MYES evaluation are sensibly motivated such that no adjustment is typically required.

The vast MYES serves as a proxy for the still larger ACS. We argue for the validity of this extrapolation on the grounds of the initial design of the MYES – it was constructed to be just such an antecedent proxy for the more comprehensive ACS. Nevertheless, our methods can soon be applied to the actual ACS database; all that is required is sufficient time such that a wealth of one-year estimates are available for comparison to published three- and five-year MYEs. Thus, in the future an analyst who wishes to violate the U.S. Census Bureau injunction – by making an average, end-year, or mid-year substitution of one-year estimates for a three- or five-year MYE – can first check whether this is consequential by utilizing relative percent differences.

A more nuanced scrutiny of the comparability question is possible, and is demonstrated through the 20 specific variables discussed in Section 5.2. This demonstrates the strength of our method – the ability to focus on a particular variable of interest without requiring additional assumptions. The story with these 20 variables is largely consistent with our overall thesis, although there are some interesting surprises as well. Overall, we find solid support that comparability is dangerous, though the exact consequence depends on many factors: period length; Numeric versus Percent; variable category and exact variable; region (county in our study); average versus middle versus end. We conclude by reinforcing the U.S. Census Bureau cautions against making such *ad hoc* comparisons.

References

[1] Alexander, C. (1998) Recent developments in the American Community Survey. *Proceedings of the Survey Research Methods Section, ASA*, 92–100.

[2] Beaghen, M. and Weidman, L. (2008) "Statistical Issues of Interpretation of the American Community Survey's One-, Three-, and Five-Year Period Estimates," U.S. Census Bureau, 2008 American Community Survey Research Memorandum Series #ACS08-R-4.

[3] Beaghen, M. and Stern, S. (2009) "Usability of the American Community Survey Estimates of the Group Quarters Population for Substate Geographies," *Proceedings of the 2009 Joint Statistical Meetings on CD-ROM* American Statistical Association, pp. 2123-2137.

[4] Beaghen, M., Weidman, L., Asiala, M., McElroy, T., and Navarro, A. (2010) *Interpretation and Use of American Community Survey Multiyear Estimates. Mimeo.*

[5] Bell, W. (1998) "Borrowing Information Over Time in Small Area Estimation: Thoughts with Reference to the American Community Survey," unpublished paper presented at the American Community Survey workshop, September 13, 1998, Michael Cohen (workshop study director), Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, National Research Council, Washington, DC.

[6] Breidt, F. (2007) "Alternatives to the Multiyear Period Estimation Strategy for the American Community Survey," Appendix C in *Using the American Community Survey: Benefits and Challenges, National Research Council, Panel on the Functionality and Usability of Data from the American Community Survey*, Constance F. Citro and Graham Kalton, editors, Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press.

[7] Citro, C. and Kalton, G. (2007) *Using the American Community Survey: Benefits and Challenges, National Research Council, Panel on the Functionality and Usability of Data from the American Community Survey*, Constance F. Citro and Graham Kalton, editors, Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press.

[8] Fay, R. (2007) "Imbedding model-assisted estimation into ACS estimation," *Proceedings of the 2007 Joint Statistical Meetings on CD-ROM,* American Statistical Association, pp. 2946–2953.

[9] Kish, L. (1981) "Using cumulated rolling samples to integrate census and survey operations of the Census Bureau," Washington, D.C., U.S. Government Printing Office.

[10] McElroy, T. (2009) Incompatibility of Trends in Multi-Year Estimates from the American Community Survey. *The Annals of Applied Statistics* **3**, 1493–1504.

[11] Navarro, A. (2009) Private Communication.

[12] Starsinic, M. and Tersine, A. (2007) "Analysis of variance estimates from American Community Survey multiyear estimates," *Proceedings of the 2007 Joint Statistical Meetings on CD-ROM,* American Statistical Association, pp. 3011–3017.

[13] Tersine, A. and Asiala, M. (2007) "Methodology for the production of American Community Survey multiyear estimates," *Proceedings of the 2007 Joint Statistical Meetings on CD-ROM,* American Statistical Association, pp. 3018–3023.

[14] Torrieri, N. (2007) America is Changing, and so is the Census: The American Community Survey," *The American Statistician* **61**, 16–21.

[15] U.S. Census Bureau (2006) Technical Paper 67. *Design and Methodology, American Community Survey.*
http://www.census.gov/acs/www/Downloads/tp67.pdf

[16] U.S. Census Bureau (2007) Multi-Year Estimates Study, "Accuracy of the Data."
http://www.census.gov/acs/www/Downloads/Accuracy of the Data MYE.pdf

A Summary Tables

Table 3: Counties' names and their ID numbers.

County name	ID
Black Hawk County, Iowa	19013
Bronx County, NewYork	36005
Broward County, Florida	12011
Calvert County, Maryland	24009
Douglas County, Nebraska	31055
Flathead County, Montana	30029
Franklin County, Ohio	39049
Hampden County, Massachusetts	25013
Jefferson County, Arkansas	05069
Lake County, Illionois	17097
Madison County, Mississippi	28089
Multnomah County, Oregon	41051
Pima County, Arizona	04019
Rockland County, NewYork	36087
San Francisco, California	06075
Schuylkill County, Pennsylvania	42107
Sevier County, Tennessee	47155
Tulare County, California	06107
Yakima County, Washington	53077

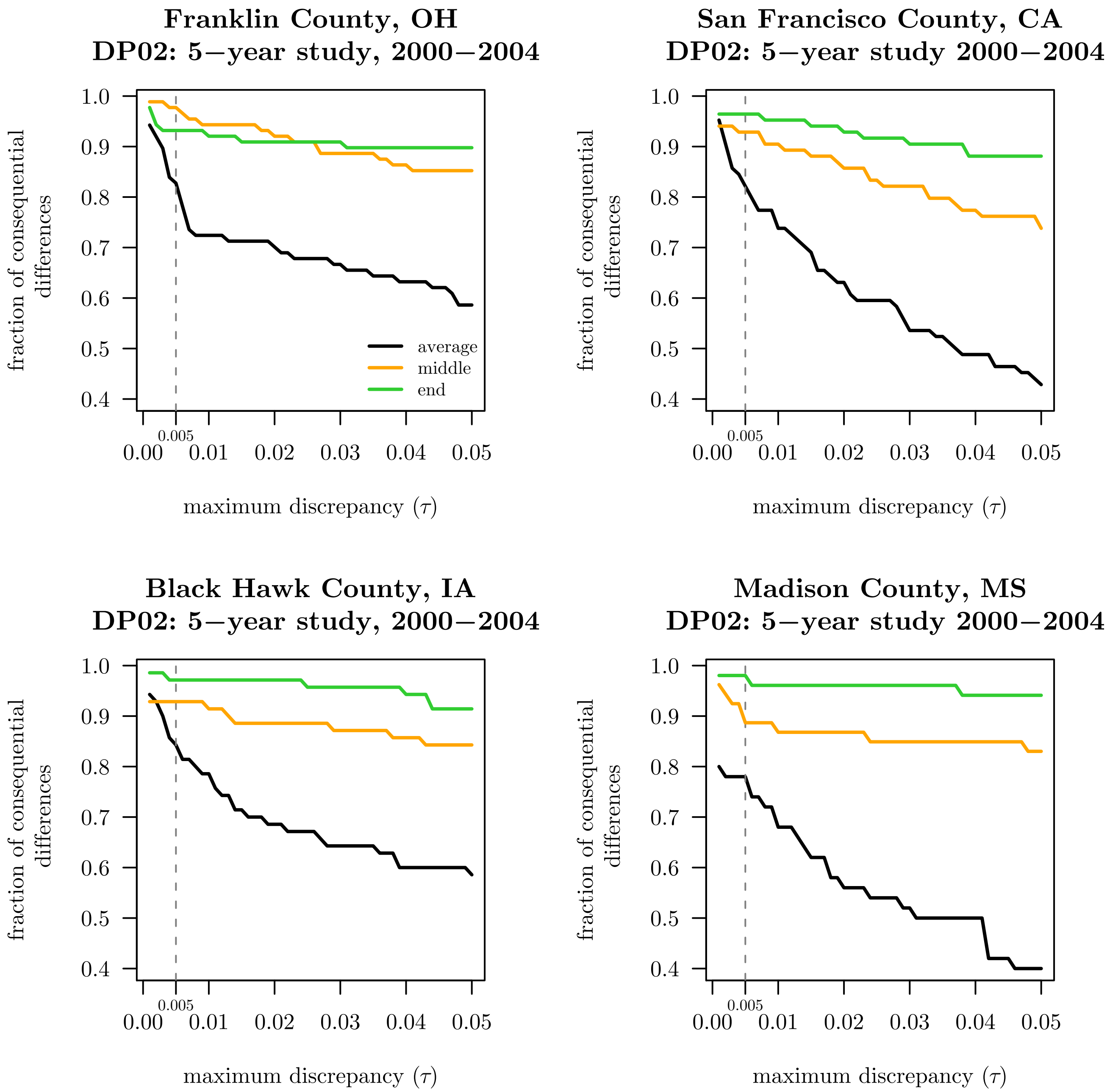

Figure 3: Fraction of consequential differences by maximum discrepancy τ

Table 4: Proportions of consequential relative percent differences (three-year studies only).

Study		$\lvert D^{(3)}_{ave,t}\rvert$		$\lvert D^{(3)}_{mid,t}\rvert$		$\lvert D^{(3)}_{end,t}\rvert$	
		Numerics	Percents	Numerics	Percents	Numerics	Percents
Table $DP01$: Demographic variables							
2000-2002	Average	0.70	0.68	0.91	0.79	0.94	0.84
	Maximum	0.84	0.82	0.97	0.92	0.99	0.98
2001-2003	Average	0.66	0.66	0.90	0.80	0.93	0.82
	Maximum	0.76	0.80	0.98	0.89	0.98	0.89
2002-2004	Average	0.67	0.65	0.90	0.82	0.95	0.83
	Maximum	0.83	0.81	0.96	0.89	1.00	0.96
2003-2005	Average	0.66	0.65	0.92	0.80	0.94	0.83
	Maximum	0.86	0.82	0.97	0.90	0.98	0.96
Table $DP02$: Social variables							
2000-2002	Average	0.46	0.54	0.90	0.79	0.92	0.83
	Maximum	0.64	0.76	0.95	0.94	0.96	0.89
2001-2003	Average	0.44	0.53	0.12	0.82	0.08	0.84
	Maximum	0.60	0.70	0.22	0.97	0.15	0.95
2002-2004	Average	0.47	0.53	0.11	0.83	0.10	0.85
	Maximum	0.71	0.67	0.23	0.92	0.17	1.00
2003-2005	Average	0.46	0.49	0.90	0.83	0.92	0.83
	Maximum	0.69	0.60	0.97	1.00	0.97	0.94
Table $DP03$: Economic variables							
2000-2002	Average	0.61	0.56	0.91	0.87	0.93	0.89
	Maximum	0.72	0.73	0.97	0.95	0.96	0.96
2001-2003	Average	0.60	0.53	0.90	0.88	0.91	0.88
	Maximum	0.79	0.71	0.96	0.97	0.96	0.95
2002-2004	Average	0.57	0.55	0.12	0.87	0.11	0.88
	Maximum	0.83	0.74	0.21	0.94	0.18	0.97
2003-2005	Average	0.60	0.55	0.92	0.89	0.89	0.87
	Maximum	0.77	0.69	0.96	0.94	0.95	0.93
Table $DP04$: Housing variables							
2000-2002	Average	0.50	0.54	0.89	0.87	0.91	0.89
	Maximum	0.69	0.69	0.94	0.92	0.96	0.94
2001-2003	Average	0.46	0.50	0.89	0.88	0.91	0.88
	Maximum	0.65	0.61	0.94	0.93	0.96	0.93
2002-2004	Average	0.49	0.50	0.89	0.88	0.91	0.89
	Maximum	0.68	0.67	0.94	0.92	0.95	0.94
2003-2005	Average	0.49	0.54	0.90	0.87	0.92	0.89
	Maximum	0.63	0.68	0.94	0.93	0.98	0.94

Table 5: Proportions of consequential relative percent differences (five-year studies only).

Study		$\lvert D^{(5)}_{ave,t}\rvert$		$\lvert D^{(5)}_{mid,t}\rvert$		$\lvert D^{(5)}_{end,t}\rvert$	
		Numerics	Percents	Numerics	Percents	Numerics	Percents
Table DP01: Demographic variables							
2000-2004	Average	0.90	0.88	0.93	0.85	0.99	0.94
	Maximum	1.00	1.00	0.98	0.92	1.00	1.00
2001-2005	Average	0.91	0.89	0.94	0.84	0.99	0.94
	Maximum	1.00	0.99	0.98	0.91	1.00	1.00
Table DP02: Social variables							
2000-2004	Average	0.84	0.83	0.93	0.86	0.98	0.93
	Maximum	1.00	1.00	0.97	0.97	1.00	0.98
2001-2005	Average	0.85	0.87	0.93	0.86	0.98	0.94
	Maximum	1.00	1.00	0.98	0.94	1.00	1.00
Table DP03: Economic variables							
2000-2004	Average	0.88	0.86	0.94	0.91	0.97	0.95
	Maximum	1.00	0.98	0.98	0.98	1.00	0.98
2001-2005	Average	0.89	0.87	0.93	0.91	0.97	0.95
	Maximum	1.00	1.00	0.98	0.97	1.00	1.00
Table DP04: Housing variables							
2000-2004	Average	0.86	0.85	0.94	0.91	0.98	0.96
	Maximum	1.00	0.99	0.98	0.96	1.00	1.00
2001-2005	Average	0.87	0.85	0.92	0.91	0.97	0.95
	Maximum	1.00	0.99	0.96	0.97	1.00	1.00

Table 6: Variables selected for further examination.

Profile line	**Variable Name**
Table $DP01$: Demographic variables	
2	Number of males
17	Median age
42	Number of Filipinos
103	Average house hold size
Table $DP02$: Social variables	
31	Fertility per 1,000 unmarried women
37	Number of grandparents who are responsible for their grandchildren
41	Veteran status/civilian veterans
59	Not a US Citizen
79	Speak English less than "very well"
Table $DP03$: Economic variables	
25	Mean travel time to work
49	Number of unpaid family workers
51	Income $\$10,000$ and under
55	Income between $\$35,000 - \$49,999$
62	Mean household income
94	% of all families in poverty
Table $DP04$: Housing variables	
4	Homeowner vacancy rate
26	Number of rooms is 5
59	House heating fuel is solar energy
72	Value of housing unit between $\$150,000 - \$199,000$
99	Selected monthly house owner costs as a percentage of household income, 35% or more

Table 7: Proportions of consequential relative percent differences for selected variables.

Profile line		$\lvert D^{(3)}_{ave,t}\rvert$	$\lvert D^{(3)}_{mid,t}\rvert$	$\lvert D^{(3)}_{end,t}\rvert$	$\lvert D^{(5)}_{ave,t}\rvert$	$\lvert D^{(5)}_{mid,t}\rvert$	$\lvert D^{(5)}_{end,t}\rvert$
Table $DP01$: Demographic variables							
2	Numeric	0.961	0.987	0.987	1	1	1
	Percent	0.776	0.513	0.684	0.974	0.632	0.868
17	Numeric	0.789	0.724	0.961	0.842	0.789	1
	Percent	-	-	-	-	-	-
42	Numeric	0.357	0.907	0.932	0.864	0.864	1
	Percent	0.381	0.884	0.932	0.909	0.864	1
103	Numeric	0.763	0.803	0.947	0.842	0.895	0.921
	Percent	-	-	-	-	-	-
Table $DP02$: Social variables							
31	Numeric	0.493	0.893	0.88	0.694	0.946	0.895
	Percent	-	-	-	-	-	-
37	Numeric	0.356	0.907	0.932	0.865	0.892	0.947
	Percent	0.315	0.867	0.932	0.703	0.919	0.947
41	Numeric	0.485	0.958	0.931	0.879	1	0.974
	Percent	-	-	-	-	-	-
59	Numeric	0.539	0.882	0.934	0.868	0.895	1
	Percent	0.368	0.855	0.868	0.921	0.921	0.974
79	Numeric	0.66	0.863	0.98	0.826	0.923	1
	Percent	0.702	0.804	0.918	0.913	0.923	0.97
Table $DP03$: Economic variables							
25	Numeric	0.461	0.855	0.829	0.868	0.842	0.921
	Percent	-	-	-	-	-	-
49	Numeric	0.263	0.868	0.855	0.763	0.921	1
	Percent	0.658	0.566	0.592	0.816	0.605	0.737
51	Numeric	0.882	0.934	0.908	1	0.868	0.974
	Percent	0.895	0.895	0.908	0.974	0.868	0.947
55	Numeric	0.658	0.908	0.934	0.947	0.974	1
	Percent	0.684	0.934	0.947	0.842	1	0.974
62	Numeric	1	0.947	0.934	1	1	0.947
	Percent	-	-	-	-	-	-
94	Numeric	0.483	0.925	0.906	0.733	0.969	0.973
	Percent	-	-	-	-	-	-

Table 7: Proportions of consequential relative percent differences for selected variables.

Table $DP04$: Housing variables							
4	Numeric	0.267	0.816	0.947	0.658	0.868	0.947
	Percent	-	-	-	-	-	-
26	Numeric	0.368	0.947	0.921	0.816	0.947	1
	Percent	0.395	0.947	0.947	0.737	0.947	0.974
59	Numeric	0.204	0.98	0.898	0.75	0.964	0.964
	Percent	0.583	0.417	0.417	0.857	0.429	0.571
72	Numeric	0.452	0.865	0.973	0.919	0.946	0.974
	Percent	0.315	0.905	0.932	0.973	0.838	0.974
99	Numeric	0.211	0.855	0.934	0.789	0.842	0.974
	Percent	0.342	0.882	0.908	0.658	0.868	0.947