NBER WORKING PAPER SERIES

EASURING LABOR EARNINGS INEQUALITY USING PUBLIC-USE MARCH CURRENT POPULATION SURVEY DAT THE VALUE OF INCLUDING VARIANCES AND CELL MEANS WHEN IMPUTING TOPCODED VALUES

Richard V. Burkhauser
Shuaizhang Feng
Jeff Larrimore

Working Paper 14458
http://www.nber.org/papers/w14458

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2008

The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the New York Census Research Data Center at Cornell University. Conclusions expressed are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau nor those of the National Bureau of Economic Research. This paper has been screened to ensure that no confidential data are disclosed. Supports for this research from the National Science Foundation (award nos. SES-0427889 SES-0322902, and SES-0339191) and the National Institute for Disability and Rehabilitation Research (H133B040013 and H133B031111) are cordially acknowledged. We thank Lisa Marie Dragoset, Ian Schmutte, Arnie Reznek, Laura Zayatz, the Cornell Census RDC Administrators, and all their U.S. Census Bureau colleagues who have helped with this project. We thank Peter Gottschalk for convincing us to undertake this effort. Burkhauser completed this paper while he was the R.I. Downing Fellow in Social Economics in the Faculty of Economics and Commerce, University of Melbourne.

Measuring Labor Earnings Inequality using Public-Use March Current Population Survey Data: The Value of Including Variances and Cell Means When Imputing Topcoded Values

Richard V. Burkhauser, Shuaizhang Feng, and Jeff Larrimore

NBER Working Paper No. 14458
October 2008
JEL No. C8,D31,J3

ABSTRACT

Using the Census Bureau's internal March Current Population Surveys (CPS) file, we construct and make available variances and cell means for all topcoded income values in the public-use version of these data. We then provide a procedure that allows researchers with access only to the public-use March CPS data to take advantage of this added information when imputing its topcoded income values. As an example of its value we show how our new procedure improves on existing imputation methods in the labor earnings inequality literature.

Richard V. Burkhauser
Cornell University
Department of Policy Analysis & Management
125 MVR Hall
Ithaca, NY 14853-4401
and NBER
rvb1@cornell.edu

Jeff Larrimore
Department of Economics
404 Uris Hall
Cornell University
Ithaca NY, 14853
jhl42@cornell.edu

Shuaizhang Feng
Princeton University
Woodrow Wilson School
Princeton, NJ 08544
shuaizhang.feng@gmail.com

Introduction

For confidentiality reasons, the Census Bureau topcodes all values above some maximum for each income source in the public-use version of the March Current Population Survey (CPS). Since 1996, the Census Bureau has provided the cell mean of all topcoded incomes reported for at least some of these income sources, including the four income sources related to labor earnings.[1]

With these cell means, researchers using the public-use CPS can match the mean values found using internal CPS data for these income sources. Most researchers interested in measuring long-term trends in earnings including years prior to 1996, however, have imputed their own topcode values to create a consistent series. A common practice is to impute topcoded earnings as a fixed multiple above the topcode point, although there is little consensus on the appropriate multiple to use. For example, Juhn, Murphy, and Pierce (1993) impute weekly wages and salaries for workers with topcoded earnings at 133% of their topcoded value using 1963–1989 public-use CPS data and ignored changes over time in the proportion of the weekly wages and salaries distribution affected by the topcode rule. Lemieux (2006) multiplies topcoded hourly wages and salaries by a factor 1.4 using the public-use CPS, while Autor, Katz, and Kearney (2008) and Katz and Murphy (1992) multiply topcoded weekly wages and salaries by a factor of 1.5.[2]

In order to further alleviate concerns about topcoding, many researchers who impute topcoded wages as a fixed multiple above the topcode point also focus on the P90/P10 ratio rather than inequality statistics such as the Gini coefficient or the Theil index which incorporate information about the full distribution of wages and salaries. However, Burkhauser, Feng, and Jenkins (2007) demonstrate that while the P90/P10 ratios for both wages and salaries and total

labor earnings are not greatly impacted by topcoding, the trends using the P90/P10 ratio are different than those found using the Gini coefficient. Thus, a more precise topcode correction method would allow for a further analysis of inequality trends using inequality statistics that encompass more of the earnings distribution than the P90/P10 ratio.

Fichtenbaum and Shahidi (1988) use a slightly more sophisticated topcode correction method, setting all values above the topcode threshold in the public-use CPS to a single value such that the distribution of annual wages and salaries above the median best fits the Pareto distribution. Bernstein and Mishel (1997) use a similar approach, but fit the entire distribution of hourly wages and salaries to the Pareto distribution instead of just the upper half of the distribution. As a result, if the distribution of wages and salaries shifts to the right, then the imputed value for topcoded wages and salaries will increase even if the topcode threshold is unchanged. Bernstein and Mishel note that such a Pareto imputation for topcoded values in public-use CPS is not perfect, as it involves imputing one value for all top-coded wages and salaries but they argue that it is better than ignoring topcodes altogether. See Angle and Tolbert (1999) for a further discussion of using a Pareto distribution to impute a single value for topcoded wage and salaries .

With access to internal CPS data used by the Census Bureau in their official estimations of income and its distribution,[3] we have constructed and disseminate cell means for the years 1976-2002 (Larrimore, Burkhauser, Feng and Zayatz, 2008).[4] With these cell means researchers who previously assigned all topcoded earnings values an estimate of the mean of all of those earnings values above the topcode threshold can now assign these individuals their mean values above the threshold based on the internal data. However, this approach is still limited because it assumes all individuals with a topcoded earnings value have the identical earnings value for that

type of earnings.

A potential remedy for the problem of assigning all individuals with topcoded earnings the same value is to assume that topcoded earnings fit the Pareto distribution. It is then possible to use a multiple imputation approach to assign earnings to topcoded individuals such that the distribution of their earnings both fit the Pareto distribution and has a mean that matches the cell means in the internal data.

While this approach is consistent with the common assumption that the earnings distribution roughly follows a Pareto distribution, we believe one can do even better than Pareto when matching earnings to those seen in the internal CPS data. In this paper, we construct and disseminate the variances of each cell for topcoded incomes in the public-use March CPS. This information can be used together with our cell means to further improve estimates at the top of the income distribution. While the Pareto distribution is convenient since it can be fitted using only the mean earnings above the topcode, it is not flexible enough to also match a known variance. Therefore, we instead assume that earnings above the topcode threshold fit a Stoppa distribution, which is a generalization of the Pareto distribution (Stoppa, 1990). Using a multiple imputation approach to fit topcoded labor earnings to the Stoppa distribution fitted around the mean and variance of topcoded labor earnings, we find that one can very closely match the distribution of labor earnings found in the internal CPS data. While we focus on the broadest definition of labor earnings, which includes income from primary earnings, wages and salaries, self-employment, and farm earnings, similar results are obtained when focusing on the narrower category of only income from wages and salaries.

In the following section we briefly explain how we calculate the cell means and variances using the internal CPS data. We then demonstrate that the log variance of labor earnings found

using a multiple imputation procedure based on public-use CPS data with our cell mean and variance data better matches the log variance using the internal CPS than other methods used in the literature. While we focus on the log variance of labor earnings, such an imputations procedure could similarly be used to obtain more accurate values of other statistics of interest such as the Gini Coefficient or could be applied to other definitions of income including earnings from wages and salaries or total household income including non-labor income. More generally, any researcher using the public-use CPS data can now use these newly available variance data to improve their imputations of topcoded incomes.

Cell Means and Variances for Topcoded Incomes

In calculating the cell mean and variance for each source of topcoded income in the public-use March CPS from 1976-2004, we attempted to remain consistent with the procedures used by the Census Bureau to provide its cell means in recent years. Since topcoding in the public-use CPS occurs for each separate income source, we calculated cell means and variances for each source of income separately—11 prior to 1988 and 24 since then. Also mirroring Census Bureau practices, our results are the weighted mean and variance of the income of all individuals above the topcode value for the income source.

While, in general, cell means and variances are reported separately for each income source, for confidentiality reasons cell means were combined when less than 5 individuals are topcoded from a source of income. In these instances, we combined 2 or more income sources to achieve a group of at least 5 individuals and then provide the mean and variance of topcoded incomes from this combined cell. This is the same procedure for handling small cells used by the Census Bureau in the cell means they have released in recent years (Census Bureau, 2007a). See Larrimore, Burkhauser, Feng, and Zayatz (2008) for a more detailed discussion of the

creation of the cell mean series and the procedures for protecting the confidentiality of respondents. For ease of presentation of the variance data, we took the square-root of the variance and provide the standard deviations instead. The means of topcoded incomes for each income source are provided in Appendix Tables 1 and 2 and the standard deviations are reported in Appendix Tables 3 and 4.

For those interested in more details about specific demographic groups, for each source of labor earnings we also provide cell means and variances divided by race, gender, and employment status. These are the same demographic groups used by the Census Bureau in the cell means they have released in recent years. For these demographic-based cell means, the cell mean is the weighted mean earnings from each of these sources of labor earnings for topcoded individuals with the specified demographic characteristics. As was the case for the standard cell means, if less than 5 individuals within a demographic group are topcoded on a source of labor earnings then for confidentiality reasons those individuals are combined with other demographic groups to ensure that at least 5 people are included in the cell mean calculation. The means of topcoded values for each source of labor earnings for each demographic group are provided in Appendix Tables 5 through 11 and their standard deviations are reported in Appendix Tables 12 through 18.

Imputation Using Both Cell Means and Variances

Using the cell means and variances reported above, it is possible to improve on previous methods of imputing topcoded earning values as a fixed multiple above the topcode threshold or of assigning all topcoded individuals the cell mean as their imputed earnings. Since these methods assign all topcoded individuals the same level of earnings income, they implicitly assume no variance in topcoded earnings. This causes calculations of the variance of earnings

for the entire population to be understated.

In order to improve estimates of topcoded earnings by using the variance and cell mean information from internal data, we assume that the labor earnings distribution fits the Stoppa distribution, which is a generalization of the Pareto distribution. The Stoppa distribution has a cumulative distribution function:

$$F(y) = [1 - \left(\frac{y}{y_0}\right)^{-\alpha}]^{\theta}$$

where y_0 is a lower bound of the distribution, in this case representing the topcode threshold, and the parameters α and θ specify the shape of the distribution. The Pareto distribution is the special case of the Stoppa distribution where θ equals 1.

While the Pareto distribution is often considered a good fit for the US income distribution (Piketty and Saez 2003), and has commonly been used to determine fixed-values to impute topcoded wages and salaries (see Fichtenbaum and Shahidi, 1988; Bishop, Chiou, and Formby, 1994; and Bernstein and Mishel, 1997), it is not flexible enough to incorporate known information about the variance of the distribution. As a result, while one could use the Pareto distribution to impute varying earnings for individuals above the topcode threshold, even when setting the mean of the imputed earnings correctly, such an imputation will generally estimate the variance incorrectly.

By using the more general Stoppa distribution, we can improve on imputations using the Pareto distribution. Because it is a two-parameter distribution, we can use the Stoppa distribution to incorporate the known information about both the mean and variance of topcoded incomes. Thus, instead of assuming that the variance of topcoded labor earnings fits the standard variance of the Pareto distribution, the Stoppa distribution allows us to ensure that the variance of topcoded labor earnings in our imputations match the variance in the internal CPS data.

Before imputing topcoded values, we must first calculate α and θ for each of the four sources of labor earnings (primary earnings, wages and salary, self-employment, and farm earnings) in each year such that the mean and variance of topcoded values based on the Stoppa distribution will match those seen in the internal CPS data.[1] This is done by solving for α and θ that simultaneously solve both of the first two moments of the Stoppa distribution:

$$m - \theta y_0 \beta\left(1 - \frac{1}{\alpha}, \theta\right) = 0$$

$$(m^2 + v) - \theta y_0^2 \beta\left(1 - \frac{2}{\alpha}, \theta\right) = 0$$

where m and v are the mean and variance of the distribution and $\beta(1 - \frac{1}{\alpha}, \theta)$ and $\beta(1 - \frac{2}{\alpha}, \theta)$ are beta functions.[2]

Because of natural shifts in the distribution above the topcode threshold, the parameter values for the Stoppa distribution differ from year to year and across labor earnings sources. However, some of the largest changes, such as the shift from 1981 to 1982, coincide with changes to the topcode threshold in the public-use data. This is to be expected since any change to the topcode threshold can dramatically alter the mean and variance of topcoded values. Because the parameter values change over time both due to natural fluctuations in the topcoded population and changes in the topcode thresholds, when imputing topcoded values we use different parameter values for each year and earnings source in order to provide the best fit to internal data. These parameter values for the four sources of income that make up labor earnings are provided in Table 1.[3]

Having determined the appropriate parameters in the Stoppa distribution, we can proceed with imputing each source of topcoded labor earnings. First, for each of the four sources of labor earnings (three sources prior to 1988), we find all its topcoded observation and then for each of

these observations we draw a value from the distribution that is implied by the fitted Stoppa distribution. For observations that are not topcoded, no imputation is required so we simply use the earnings as reported in the public-use CPS data. Second, we sum the four sources of labor earnings, to calculate the total labor earnings in our distribution for all individuals in each year. We estimate our inequality statistics using this distribution which includes imputed labor earnings for topcoded observations and observed labor earnings for non-topcoded observations. Third, we repeat steps 1 and 2 one hundred times, and finally, we average the one hundred estimates in each year to generate the reported inequality statistics.

Comparing the log variance of labor earnings using different topcode corrections

In Figure 1, we compare the trends in the log variance of labor earnings in the public-use March CPS using our multiply imputed data based on the Stoppa distribution (Public-Stoppa) to those obtained using four other methods of correcting for censoring in the public-use CPS data. These methods include ignoring cell means and assigning the topcode value as the labor earnings for all topcoded individuals (Public-NoCM), assigning a value equal to 1.4 times the topcode threshold to topcoded individuals (Public-Multiple),[8] assigning the cell mean value as the earnings for all topcoded individuals (Public-CM), and using our multiple imputation approach but fitting topcoded values to the Pareto distribution instead of the Stoppa distribution (Public-Pareto). Finally, we also include the log variance of labor earnings values found in the actual internal data (Internal).

We find that in all years except 1985 (the only year the public-use top code values were identical to the censoring threshold of the internal CPS data) the Public-NoCM series understates the log variance of labor earnings in the internal CPS compared to each of the other series. This should not come as a surprise since this series will suppress the values of top labor earners, thus

compressing the labor earnings distribution and reducing the labor earnings variance.

The log variance found in the Public-Multiple series is higher than that found in the Public-NoCM series because it allows for greater labor earnings at the top of the labor earnings distribution. Prior to 1994, this series does a reasonable job of approximating the log variance of labor earnings seen in the internal data. However, after 1994 when the Census increased the censoring levels in their internal data the Public-Multiple series greatly understates the amount of labor earnings held by top labor earners and as a result also understates the log variance of labor earnings. Using this procedure to match the level of labor earnings received by topcoded individuals in the internal data would require imputing topcoded sources of labor earnings at approximately twice the topcode threshold instead of 1.4 times the threshold. However, even with a more accurate multiple for estimating topcoded sources of labor earnings, this method would still understate the log variance of labor earnings because there is no variance in the imputed labor earnings for topcoded individuals.

Using cell means to impute topcoded values further improves the estimates of the log variance of labor earnings since, unlike the Public-NoCM and Public-Multiple series, the Public-CM series does not assign topcoded individuals an arbitrary level of earnings. It assigns them the actual mean labor earnings seen in the internal CPS data. However, as was the case with the Public-Multiple series, even with the most accurate estimate of the level of earnings received by topcoded individuals the log variance of labor earnings will be understated compared to the actual log variance seen in the internal CPS data because it assumes that all topcoded individuals have the same earnings.

To further improve approximations of the log variance found in the internal CPS data, it is necessary to drop the restrictive assumption that all individuals with a topcode for some source

of earnings have the same earnings values for that source of labor earnings. Prior to the availability of cell means and variances for earnings above the topcode threshold for each sources of labor earnings, the best approximation, not assuming that all topcoded individuals have the same value, was the Public-Pareto series. This series assumes topcoded earnings values fit the Pareto distribution fitted so that the mean of the distribution matches the cell mean of topcoded income sources. However, since the variance of the Pareto distribution is greater than the observed variance of topcoded labor earnings, we find that this approach greatly overstates the log variance of labor earnings in the internal CPS data. In fact, for many years the Public-Pareto series misses the log variance of labor earnings in the internal data by even more than the Public-CM series.

The Public-Stoppa series, which uses our multiple imputation approach to fit topcoded values for each sources of labor earnings to the Stoppa distribution with a mean and variance for topcoded labor earnings that matches the mean and variance of topcoded labor earnings in the internal CPS data, provides a much better fit to the log variance of labor earnings seen in the internal CPS data. Since this series does not assume that all individuals topcoded for a given source of labor earnings have the same earnings from that source, it accurately finds a higher variance than each series that make this assumption. And since it fits the variance of topcoded labor earnings to the actual variance in the internal CPS data, it provides a much better fit to the true distribution of labor earnings above the topcode in the internal CPS data than the Public-Pareto series which relies on the assumption that the variance of topcoded earnings matches the variance of the Pareto distribution.[9]

Conclusions

Topcoding in the public-use CPS data has limited the ability of researchers to incorporate

high earners in estimates of the US earnings inequality. This problem is exacerbated by non-systematic changes in topcoding over time, causing inaccuracies in trends observed based on this data. Using internal CPS data, we have created a series of cell means and variances for topcoded income which can be used in conjunction with public-use CPS data to generate more accurate estimates of US income levels and trends. When using the cell means and variance to impute topcoded labor earnings in the public-use CPS we found that the log variance of labor earnings is greater than is observed when ignoring topcodes or when using cell means alone and less than that found assuming a Pareto distribution. It much more closely mirrors the actual log variance of labor earnings using the internal CPS data than these other approximations.

Because some censoring exists in the internal March CPS data, results based on these statistics will still omit the very top earners whose labor earnings are above both the public-use topcode and the internal thresholds. For researchers interested in these earners, who are generally confined well within the top 1% of the income distribution, it is necessary to make additional out-of-sample predictions about the distribution of labor earnings above the internal censoring points (for an example of such an out-of-sample estimation, see Burkhauser, Feng, Jenkins and Larrimore, 2008). For researchers limited to using the public-use CPS, who wish to avoid the additional distributional assumptions required for such out-of-sample predictions, using variance and cell means of topcoded incomes that we provide will yield the best estimates of the US earnings and income distribution that can be made without making out of sample predictions on incomes above the internal censoring points.

References

Angle, John and Charles Tolbert, *Topcodes and the Great U-Turn in Nonmetro/Metro Wage and Salary Income Inequality. Staff Report (#9904).* (Washington, DC: Economic Research Service, US Department of Agriculture, 1999)

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics* (2008), 300-323.

Bernstein, Jared and Lawrence Mishel, "Has wage inequality stopped growing?" *Monthly Labor Review* (Dec. 1997), 3-16.

Bishop, John A., Jong-Rong Chiou, and John P. Formby, "Truncation Bias and the Ordinal Evaluation of Income Inequality," *Journal of Business and Economic Statistics* 12 (1994), 123–127.

Burkhauser, Richard V., Shuaizhang Feng, and Stephen Jenkins, "Using the P90/P10 Ratio to Measure US Inequality Trends with Current Population Survey Data: A View from Inside the Census Bureau Vaults," Center of Economic Studies Working Paper No. CES-WP-07-17

Burkhauser, Richard V., Shuaizhang Feng, Stephen Jenkins, and Jeff Larrimore, "Trends in United States Income Inequality Using the March Current Population Survey: The Importance of Controlling for Censoring," NBER Working Paper No. w14247 (2008).

Fichtenbaum, Rudy and Hushang Shahidi, "Truncation Bias and the Measurement of Income Inequality," *Journal of Business and Economic Statistics* 6 (1998), 335–337.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101 (1993), 410–442.

Katz, Lawrence F. and Kevin M. Murphy. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics* 107 (Feb.1992), 35-78.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *The American Economic Review* 96 (2006), 461–498.

Larrimore, Jeff, Richard V. Burkhauser, Shuaizhang Feng and Laura Zayats, "Consistent Cell Means for Topcoded Incomes in the Public Use March CPS (1975-2004)," *Journal of Economic and Social Measurement* (Forthcoming).

Piketty, Thomas and Emmanuel Saez, "Income Inequality in the United States, 1913–1998," *Quarterly Journal of Economics* 118 (2003), 1–39.

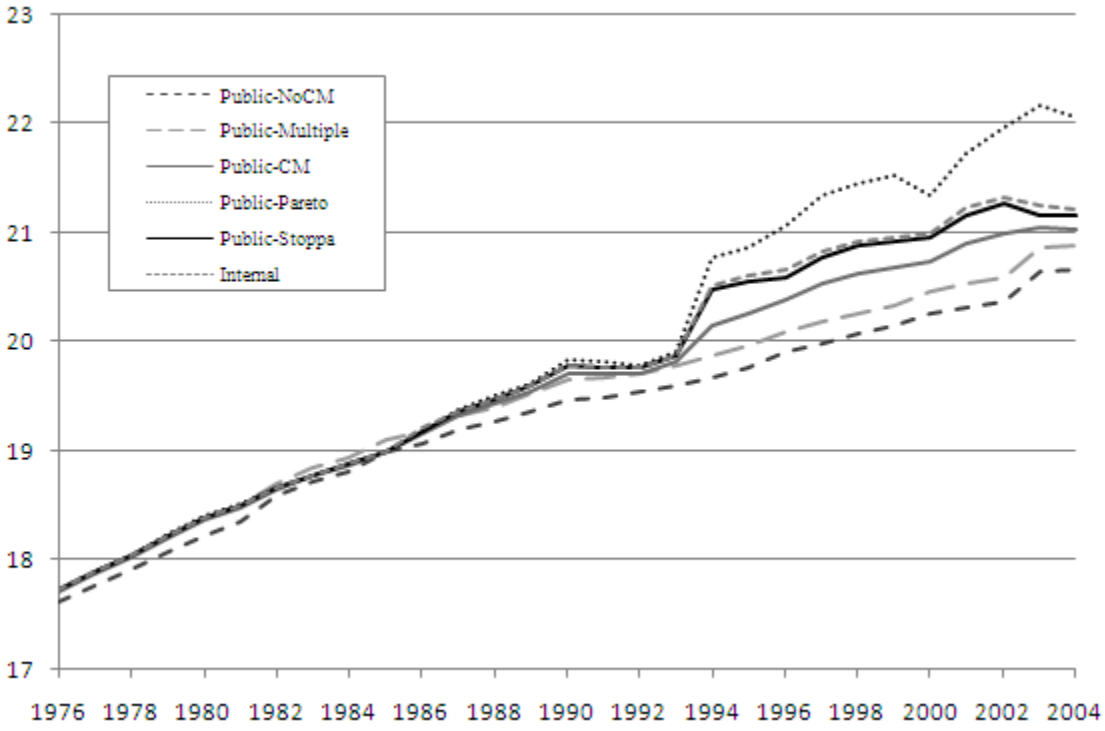
Stoppa, Gabriele, 1990. "A new model for income size distribution" (pp 33-41), in Camilo Dagum and Michele Zenga (Eds.), *Income and Wealth Distribution, Inequality and*

Poverty: Proceedings of the 2nd International Conference on Income Distribution by Size: Generation, Distribution, Measurement and Application. (New York, Berlin, London, and Tokyo: Springer, 1990)

U.S. Census Bureau, *Current Population Survey Annual Social and Economic Supplement Technical Documentation*, (Washington, DC: Government Printing Office, 2007a)

U.S. Census Bureau, *Income, Poverty, and Health Insurance Coverage in the United States: 2006. Current Population Reports, Consumer Income* (Washington, DC: Government Printing Office, 2007b)

Figure 1: Log Variance of Labor Earnings Income in the March CPS (1976-2004)



Source: Author's calculations using Public-use and Internal March CPS data

Table 1: Parameters for Stoppa Imputation

	Primary Earnings		Wages and Salary		Self Employment		Farm	
	α	θ	α	θ	α	θ	α	θ
1976	N/A[A]	N/A[A]	5.33854	1.52599	5.29697	1.99269	4.30873	1.29109
1977	N/A[A]	N/A[A]	4.97897	1.24502	5.11376	1.61624	4.16415	0.67046
1978	N/A[A]	N/A[A]	5.16776	1.50451	4.75058	1.39774	4.31424	0.61661
1979	N/A[A]	N/A[A]	5.19448	1.65199	5.02041	1.73423	4.86191	1.02943
1980	N/A[A]	N/A[A]	5.27078	1.88653	5.1079	2.12236	5.41444	1.44821
1981	N/A[A]	N/A[A]	5.20448	1.63454	5.07139	2.26725	4.97354	0.86369
1982	N/A[A]	N/A[A]	11.7997	5.01319	11.4205	4.01987	15.5396	10.2857
1983	N/A[A]	N/A[A]	10.9317	3.15905	10.1800	2.46249	9.63002	0.78819
1984	N/A[A]	N/A[A]	11.0130	3.61926	11.8257	5.25631	8.15904	0.69895
1985	N/A[A]	N/A[A]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1986	N/A[A]	N/A[A]	4.37235	1.39692	3.96431	0.84829	N/A[B]	N/A[B]
1987	N/A[A]	N/A[A]	4.27739	1.50169	4.35652	1.28696	6.33261	1.31005
1988	3.81260	1.53148	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1989	3.78297	1.40645	N/A[B]	N/A[B]	3.56481	0.570547	N/A[B]	N/A[B]
1990	3.83321	1.87518	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1991	3.89204	1.78859	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1992	3.93927	1.32414	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1993	4.04503	1.74913	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]	N/A[B]
1994	2.54640	1.64920	3.74483	1.35999	3.49265	1.51830	3.49265	1.51830
1995	2.60114	1.62852	3.61675	3.02757	2.61053	7.50127	2.61053	7.50127
1996	2.98294	3.11612	2.11152	2.25299	2.25118	4.24653	4.14202	7.12770
1997	2.79848	3.12629	2.15788	2.19959	2.29437	5.41491	5.55024	5.03893
1998	2.80358	3.12021	2.22761	2.04733	3.00114	5.71152	3.03641	16.9028
1999	2.83666	3.20251	2.27154	1.65811	2.29238	3.84018	3.09512	5.90022
2000	2.86963	2.41499	2.22923	1.51122	2.35169	2.58739	5.00936	46.5072
2001	2.76304	2.97651	2.56739	2.08882	2.29874	2.93036	2.16415	17.7604
2002	2.80519	3.10233	2.27061	1.85128	2.45704	3.11624	2.16334	20.4917
2003	3.05921	3.33213	2.28513	1.82723	2.66334	6.47645	2.20494	5.96019
2004	3.13605	3.33472	2.22761	2.04733	3.09264	4.44279	2.19971	3.84147

[A] – Primary Earnings were not reported as a separate category prior to 1988. All labor earnings were allocated separately to wage, self-employment, or farm-earnings.
[B] – Indicates that the public-use topcode was identical to the internal topcode, so the value in the public-use data matches the value in the internal data without the use of the imputations procedure.
Source: Authors Calculations using the cell-means and standard deviations from appendix tables 1-4.

Appendix Table 1: Cell Means of Topcoded Incomes (1976-1987)

Year	Wages and Salaries (I51A)	Self Employment (I51B)	Farm (I51C)	Social Security (I52A)	Supplemental Security (I52B)	Public Assistance (I53A)	Interest (I53B)	Dividends/ Rentals (I53C)	Veterans and Workers Comp (I53D)	Retirement (I53E)	Other (I53F)
1976	65193	68022	67970	N/A	N/A	N/A	73035[A]	74448	N/A	73035[A]	N/A
1977	64562	66620	61813	9999[B]	N/A	N/A	71651[A]	71651[A]	N/A	71651[A]	71651[A]
1978	65687	66665	60590	9999[B]	5999[B]	N/A	62534[A]	74304	N/A	N/A	62534[A]
1979	66514	67764	63208	9999[B]	5999[B]	N/A	66269[A]	70785	N/A	88167	66269[A]
1980	67561	69583	64447	12893	5999[B]	N/A	71163	74089	61135[A]	63786	61135[A]
1981	66367	70528	61356	11589	5999[B]	N/A	72861	71240	33808	70004[A]	70004[A]
1982	91534	90562	91015	19999[B]	5999[B]	N/A	70242[A]	93788	70242[A]	81636	70242[A]
1983	89595	88947	82381	N/A	5999[B]	19999[B]	97565[A]	92724	42207	91843	97565[A]
1984	90447	91829	83154	19999[B]	6520	N/A	94024[C]	87201	N/A	92638	89513
1985	99999[B]	99999[B]	99999[B]	19999[B]	9999[B]	N/A	99999[B]	99999[B]	43336	99999[B]	99999[B]
1986	137113	129996	N/A	19999[B]	N/A	N/A	99999[B]	99999[B]	41302	99999[B]	99999[B]
1987	140026	135346	122398	19999[B]	9999[B]	N/A	99999[B]	99999[B]	48662	99999[B]	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
[C] - Interest income in 1984 does not properly match between the internal and the public data. This cell mean is based on the 25 individuals with interest income at or above $75,000 topcode threshold in the internal data, not just the 11 people who are listed as topcoded for interest income in the public data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Note: In parentheses below each variable name is the mnemonic for the income source from the March Current Population Survey Technical Documentation.

Source: Author's calculations using internal March CPS data

Appendix Table 2: Cell-Means of Topcoded Incomes (1988-2004)

Year	Primary Earnings (ERN_VAL)	Wages and Salaries (WS_VAL)	Self Employment (SE_VAL)	Farm (FRM_VAL)	Social Security (SS_VAL)	Supplemental Security (SSI_VAL)	Public Assistance (PAW_VAL)	Interest (INT_VAL)	Dividends (DIV_VAL)	Rental (RNT_VAL)	Alimony (ALM_VAL)	Child Support (CSP_VAL)
1988	147389	99999[B]	99999[B]	99999[B]	29999[B]	9999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	N/A
1989	145322	N/A	125624	99999[B]	N/A	9999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	N/A
1990	153438	99999[B]	99999[B]	N/A	29999[B]	9999[B]	19999[B]	99999[B]	99999[B]	99999[B]	99999[B]	N/A
1991	150805	N/A	99999[B]	N/A	29999[B]	9999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	N/A
1992	141202	99999[B]	99999[B]	N/A	29999[B]	9999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	N/A
1993	147511	99999[B]	99999[B]	N/A	29999[B]	9999[B]	19999[B]	99999[B]	99999[B]	99999[B]	N/A	N/A
1994	192210	144956	153328[A]	153328[A]	49999[B]	14933	24999[B]	99999[B]	99999[B]	99999[B]	N/A	N/A
1995	188180	177066	319060[A]	319060[A]	50000[B]	14154	24999[B]	99999[B]	99999[B]	99999[B]	99999[B]	N/A
1996	308691	65394	125234	49392	50000[B]	25000[B]	25000[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]
1997	326907	63032	134769	38782	50000[B]	25000[B]	25000[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]
1998	326155	61078	98663	86012	50000[B]	25000[B]	25000[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]
1999	325365	53368	117226	60370	50000[B]	25000[B]	N/A	60820	36877	57453	93114	26592
2000	295041	52527	97311	62726	50000[B]	25000[B]	25000[B]	63005	36962	55220	54009	23918
2001	325563	51599	104848	157047	50000[B]	25000[B]	25000[B]	61107	38907	54644	62078	27947
2002	325382	55636	99362	167762	50000[B]	25000[B]	25000[B]	64853	38962	57416	63554	25657
2003	411090	77002	147370	93257	50000[B]	25000[B]	25000[B]	50186	33581	72409	54886	25714
2004	402884	82643	110899	77543	50000[B]	25000[B]	N/A	51372	39987	74636	74690	28201

Appendix Table 2 (Continued)

Year	Unemployment (UC_VAL)	Workers Comp (WC_VAL)	Veterans (VET_VAL)	Retirement 1st Source (RET_VAL1)	Retirement 2nd Source (RET_VAL2)	Survivors 1st Source (SUR_VAL1)	Survivors 2nd Source (SUR_VAL2)	Disability 1st Source (DIS_VAL1)	Disability 2nd Source (DIS_VAL2)	Education Assistance (ED_VAL)	Financial Assistance (FIN_VAL)	Other (OI_VAL)
1988	N/A	N/A	29999[B]	99999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	N/A	N/A	N/A
1989	N/A	N/A	29999[B]	99999[B]	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A	99999[B]
1990	N/A	99999[B]	29999[B]	99999[B]	N/A	99999[B]	N/A	99999[B]	N/A	N/A	N/A	99999[B]
1991	N/A	99999[B]	29999[B]	99999[B]	N/A	99999[B]	99999[B]	N/A	N/A	N/A	N/A	N/A
1992	N/A	N/A	29999[B]	99999[B]	99999[B]	99999[B]	N/A	99999[B]	N/A	99999[B]	N/A	99999[B]
1993	N/A	N/A	29999[B]	99999[B]	99999[B]	99999[B]	N/A	N/A	N/A	99999[B]	99999[B]	99999[B]
1994	N/A	N/A	N/A	99999[B]	N/A	99999[B]	N/A	99999[B]	N/A	N/A	N/A	N/A
1995	99999[B]	99999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	99999[B]	N/A	N/A	99999[B]	99999[B]
1996	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]	99999[B]	N/A	N/A	99999[B]	99999[B]
1997	N/A	N/A	99999[B]	99999[B]	N/A	99999[B]	N/A	99999[B]	N/A	99999[B]	99999[B]	99999[B]
1998	99999[B]	99999[B]	N/A	99999[B]	99999[B]	99999[B]	N/A	99999[B]	N/A	N/A	N/A	99999[B]
1999	N/A	99999[B]	99999[B]	62277	49551	88985[C]	88985[C]	55327	54856	27809	40717	44265
2000	99999[B]	99999[B]	N/A	63210	62139	83427[C]	83427[C]	73486[D]	73486[D]	25585	54559	50217
2001	99999[B]	99999[B]	99999[B]	64547	62457	82077	84408	60657[D]	60657[D]	31902	48059	40865
2002	99999[B]	N/A	99999[B]	64501	65080	90937[C]	90937[C]	52859	N/A	34876	46155	47179
2003	99999[B]	99999[B]	99999[B]	63889	81934	86528[C]	86528[C]	66563	N/A	31142	60847	58193
2004	99999[B]	99999[B]	99999[B]	62401	61613	82330[C]	82330[C]	57374	N/A	30986	51555	47050

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data
[C] - Indicates that there are not enough individuals topcoded for INCSI2 to report a cell mean in this year, so the cell mean reported is a combined cell mean for INCSI1 and INCSI2
[D] - Indicates that there are not enough individuals topcoded for INCDS2 to report a cell mean in this year, so the cell mean reported is a combined cell mean for INCDS1 and INCDS2
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Note: In parentheses below each variable name is the mnemonic for the income source from the March Current Population Survey Technical Documentation.

Source: Author's calculations using internal March CPS data

Appendix Table 3: Standard Deviations of Topcoded Incomes (1976-1987)

Year	Wages and Salaries (I51A)	Self Employment (I51B)	Farm (I51C)	Social Security (I52A)	Supplemental Security (I52B)	Public Assistance (I53A)	Interest (I53B)	Dividends/ Rentals (I53C)	Veterans and Workers Comp (I53D)	Retirement (I53E)	Other (I53F)
1976	16434	17845	22387	N/A	N/A	N/A	24943[A]	22780	N/A	24943[A]	N/A
1977	17348	17881	19109	0[B]	N/A	N/A	19398[A]	19398	N/A	19398[A]	19398[A]
1978	17240	19381	17461	0[B]	0[B]	N/A	11549[A]	19259	N/A	N/A	11549[A]
1979	17537	18783	17027	0[B]	0[B]	N/A	17698[A]	20694	N/A	9233	17698[A]
1980	17738	19249	15848	2603	0[B]	N/A	23223	18611	18844[A]	16643	18844[A]
1981	17433	19812	15544	1693	0[B]	N/A	21872	19530	6545	18942	18942
1982	10135	10264	7704	0[B]	0[B]	N/A	25970[A]	8130	25970[A]	4171	25970[A]
1983	10479	11021	9163	N/A	0[B]	0[B]	3699[A]	9056	9084	12283	3699[A]
1984	10595	10167	10900	0[B]	702	N/A	9875	9567	3905	10375	7880
1985	0[B]	0[B]	0[B]	0[B]	0[B]	N/A	0[B]	0[B]	13903	0[B]	0[B]
1986	44692	45300	N/A	0[B]	N/A	N/A	0[B]	0[B]	5985	0[B]	0[B]
1987	47505	43867	24402	0[B]	0[B]	N/A	0[B]	0[B]	17251	0[B]	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for a combined group with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
[C] - Interest income in 1984 does not properly match between the internal and the public data. This standard deviation is based on the 25 individuals with interest income at or above $75,000 topcode threshold in the internal data, not just the 11 people who are listed as topcoded for interest income in the public data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Note: In parentheses below each variable name is the mnemonic for the income source from the March Current Population Survey Technical Documentation.

Source: Author's calculations using internal March CPS data

Appendix Table 4: Standard Deviations of Topcoded Incomes (1988-2004)

Year	Primary Earnings (ERN_VAL)	Wages and Salaries (WS_VAL)	Self Employment (SE_VAL)	Farm (FRM_VAL)	Social Security (SS_VAL)	Supplemental Security (SSI_VAL)	Public Assistance (PAW_VAL)	Interest (INT_VAL)	Dividends (DIV_VAL)	Rental (RNT_VAL)	Alimony (ALM_VAL)	Child Support (CSP_VAL)
1988	59388	0[B]	0[B]	0[B]	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	N/A
1989	58674	N/A	47822	0[B]	N/A	0[B]	N/A	0[B]	0[B]	0[B]	N/A	N/A
1990	62627	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	N/A
1991	59873	N/A	0[B]	N/A	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	N/A
1992	53196	0[B]	0[B]	N/A	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	N/A
1993	55179	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	N/A	N/A
1994	172179	59210	70945[A]	70945[A]	0[B]	5611	0[B]	0[B]	0[B]	0[B]	N/A	N/A
1995	158939	82046	287365[A]	287365[A]	0[B]	5253	0[B]	0[B]	0[B]	0[B]	0[B]	N/A
1996	200524	144394	184316	19395	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]
1997	242465	115689	183078	10218	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]
1998	240919	107775	64927	56483	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]
1999	234624	71547	158137	37515	0[B]	0[B]	N/A	22672	24609	25834	9888	10154
2000	204198	76718	116278	19564	0[B]	0[B]	0[B]	21907	23250	24040	12187	12994
2001	247774	46044	138178	297808	0[B]	0[B]	0[B]	22332	27184	25141	18209	21058
2002	239975	75367	103144	319287	0[B]	0[B]	0[B]	23307	26328	27941	21505	14308
2003	255252	101225	125842	154709	0[B]	0[B]	0[B]	24620	23187	21562	6941	15075
2004	238886	123992	68349	122800	0[B]	0[B]	N/A	24731	27018	21343	21296	17908

Appendix Table 4 (Continued):

Year	Unemployment (UC_VAL)	Workers Comp (WC_VAL)	Veterans (VET_VAL)	Retirement 1st Source (RET_VAL1)	Retirement 2nd Source (RET_VAL2)	Survivors 1st Source (SUR_VAL1)	Survivors 2nd Source (SUR_VAL2)	Disability 1st Source (DIS_VAL1)	Disability 2nd Source (DIS_VAL2)	Education Assistance (ED_VAL)	Financial Assistance (FIN_VAL)	Other (OI_VAL)
1988	N/A	N/A	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	N/A	N/A	N/A
1989	N/A	N/A	0[B]	0[B]	N/A	0[B]	N/A	N/A	N/A	N/A	N/A	0[B]
1990	N/A	0[B]	0[B]	0[B]	N/A	0[B]	N/A	0[B]	N/A	N/A	N/A	0[B]
1991	N/A	0[B]	0[B]	0[B]	N/A	0[B]	0[B]	N/A	N/A	N/A	N/A	N/A
1992	N/A	N/A	0[B]	0[B]	0[B]	0[B]	N/A	0[B]	N/A	0[B]	N/A	0[B]
1993	N/A	N/A	0[B]	0[B]	0[B]	0[B]	N/A	N/A	N/A	0[B]	0[B]	0[B]
1994	N/A	N/A	N/A	0[B]	N/A	0[B]	N/A	0[B]	N/A	N/A	N/A	N/A
1995	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	0[B]	N/A	N/A	0[B]	0[B]
1996	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	0[B]	N/A	N/A	0[B]	0[B]
1997	N/A	N/A	0[B]	0[B]	N/A	0[B]	N/A	0[B]	N/A	0[B]	0[B]	0[B]
1998	0[B]	0[B]	N/A	0[B]	0[B]	0[B]	N/A	0[B]	N/A	N/A	N/A	0[B]
1999	N/A	0.001	0[B]	18418	4114	15425[C]	15425[C]	21512	25912	7722	12005	23170
2000	0[B]	0.001	N/A	16675	9087	83427[C]	83427[C]	27107[D]	27107[D]	5557	10794	26905
2001	0[B]	0.001	0[B]	18172	15965	16975	14489	24914[D]	24914[D]	12806	17013	19171
2002	0[B]	N/A	0[B]	18900	19132	16027[C]	16027[C]	21146	N/A	13280	17014	23852
2003	0[B]	0[B]	0[B]	17867	20797	18467[C]	18467[C]	21295	N/A	9113	30086	28357
2004	0[B]	0[B]	0[B]	15731	11527	19102[C]	19102[C]	23977	N/A	10295	17227	20809

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is a combined standard deviation with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data
[C] - Indicates that there are not enough individuals topcoded for INCSI2 to report a cell mean in this year, so the standard deviation reported is a combined standard deviation for INCSI1 and INCSI2
[D] - Indicates that there are not enough individuals topcoded for INCDS2 to report a cell mean in this year, so the standard deviation reported is a combined standard deviation for INCDS1 and INCDS2
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Note: In parentheses below each variable name is the mnemonic for the income source from the March Current Population Survey Technical Documentation.

Source: Author's calculations using internal March CPS data

Appendix Table 5: Cell Means of Topcoded Earnings for Wages and Salaries (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	65013	N/A	60612[A]	60612[A]	N/A	N/A	68598	N/A	N/A	N/A	N/A	N/A
1977	64197	N/A	62351[A]	63207	N/A	N/A	70337	62351[A]	N/A	N/A	N/A	N/A
1978	65436	70047[A]	70047[A]	59625	70047[A]	N/A	67600	N/A	70047[A]	70047[A]	N/A	N/A
1979	66128	64292[A]	73370	63549	64292[A]	N/A	72495	N/A	64292[A]	65016	N/A	N/A
1980	67737	61925	64371	62874	72125[A]	72125[A]	68818	N/A	70963	72125[A]	N/A	N/A
1981	66210	58990	68661	60852	55631[A]	55631[A]	70415	65397	55631[A]	85119	N/A	55631[A]
1982	91610	N/A	90491[A]	83489	N/A	N/A	94607	90491[A]	N/A	N/A	N/A	N/A
1983	89485	87647[A]	96915	92340	87647[A]	N/A	88228	N/A	87647[A]	87647[A]	N/A	N/A
1984	90220	N/A	92530[A]	88528	N/A	N/A	95586	N/A	N/A	92530[A]	N/A	92530[A]
1985	99999	99999	99999	99999	N/A	N/A	99999	N/A	N/A	99999	N/A	N/A
1986	136613	170804[A]	124324	133348	N/A	N/A	137028	N/A	N/A	170804[A]	N/A	N/A
1987	140359	119934	150042[A]	125434	169047	N/A	137893	150042[A]	N/A	150042[A]	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 6: Cell Means of Topcoded Earnings for Self-Employment Earnings (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	69286	N/A	62217[A]	N/A	N/A	N/A	62217[A]	N/A	62217[A]	N/A	N/A	N/A
1977	65763	79290[A]	79290[A]	79290[A]	N/A	N/A	68452	N/A	N/A	N/A	N/A	N/A
1978	67885	N/A	56602[A]	56602[A]	N/A	N/A	62707	56602[A]	N/A	56602[A]	N/A	N/A
1979	68168	56976[A]	56976[A]	53503	N/A	N/A	71237	N/A	N/A	56976[A]	N/A	N/A
1980	69954	72753 [A]	55589	72753[A]	N/A	N/A	70074	N/A	72753[A]	72753[A]	N/A	N/A
1981	71645	56602 [A]	73083	81331	N/A	N/A	63109	N/A	56602[A]	56602 [A]	N/A	N/A
1982	89700	N/A	95061[A]	95061[A]	N/A	N/A	92898	N/A	95061[A]	95061[A]	N/A	N/A
1983	88987	90964[A]	90964[A]	90608	N/A	90964 [A]	87585	N/A	N/A	N/A	N/A	N/A
1984	92506	86400[A]	86400[A]	89529	N/A	N/A	89588	N/A	N/A	86400[A]	N/A	N/A
1985	99999	N/A	99999	99999	N/A	N/A	99999	N/A	99999	99999	N/A	N/A
1986	136144	108836[A]	N/A	108836[A]	N/A	108836[A]	106879	N/A	N/A	108836[A]	N/A	108836[A]
1987	130751	N/A	170968[A]	170968[A]	N/A	N/A	144319	N/A	170968[A]	170968[A]	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 7: Cell Means of Topcoded Earnings for Farm Earnings (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	67970[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	67970[A]	N/A	N/A
1977	61813[A]	N/A	61813[A]	N/A	N/A	N/A	N/A	N/A	N/A	61813[A]	N/A	N/A
1978	60590	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1979	63208[A]	N/A	N/A	N/A	N/A	N/A	63208[A]	N/A	N/A	N/A	N/A	N/A
1980	64447[A]	N/A	N/A	N/A	N/A	N/A	64447[A]	N/A	N/A	N/A	N/A	N/A
1981	61356[A]	N/A	61356[A]	N/A	N/A	N/A	N/A	N/A	N/A	61356[A]	N/A	N/A
1982	91015[A]	N/A	N/A	N/A	N/A	N/A	91015[A]	N/A	N/A	N/A	N/A	N/A
1983	82381	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1984	83154[A]	N/A	N/A	N/A	N/A	N/A	83154[A]	N/A	N/A	N/A	N/A	N/A
1985	99999	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1986	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1987	122398[A]	N/A	N/A	N/A	N/A	N/A	122398[A]	N/A	N/A	N/A	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 8: Cell Means of Topcoded Earnings for Primary Labor Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	148852	136582	159300	151838	153098[A]	153098[A]	124539	153098[A]	124581	143082	N/A	153098[A]
1989	143204	138971	154412	152647	137250[A]	137250[A]	160834	137250[A]	N/A	125994	137250[A]	N/A
1990	153067	159309	153072	143812	124782[A]	124782[A]	173779	124782[A]	124782[A]	115220	N/A	N/A
1991	151763	144161	135010	153090	132453	132230[A]	148513	160432	132230[A]	127641	N/A	N/A
1992	142991	133707	136560	131061	121099[A]	N/A	140743	N/A	121099[A]	128059	N/A	121099[A]
1993	148241	144800	143657	149557	114123[A]	114123[A]	141543	114123[A]	N/A	139295	N/A	114123[A]
1994	188027	232995	205449	215571	273701	159042	188154	179478[A]	179478[A]	202593	179478[A]	N/A
1995	187347	180854	179894	191029	160143	212792	173157	458721[A]	458721[A]	154367	458721[A]	N/A
1996	302536	464791	257394	283521	N/A	404595[A]	268563	404595[A]	404595[A]	576398	N/A	404595[A]
1997	318985	391150	384160	357895	454812[A]	454812[A]	325794	454812[A]	454812[A]	222550	454812[A]	N/A
1998	330658	204326	309943	306469	267650[A]	394555	330092	267650[A]	267650[A]	442032	N/A	267650[A]
1999	306732	266285	419044	402202	492661[A]	367155	348516	N/A	492661[A]	390509	492661[A]	492661[A]
2000	300974	257525	362315	256384	244810[A]	333565	284124	244810[A]	244810[A]	284141	244810[A]	N/A
2001	335049	307007	281859	288962	337247[A]	312718	321704	N/A	337247[A]	195780	337247[A]	N/A
2002	320719	326982	331937	361315	477572	331013	319740	432873[A]	432873[A]	270371	N/A	432873[A]
2003	390823	443506	562912	480608	336974	595430	487725	N/A	N/A	343894	N/A	N/A
2004	404467	360069	427630	390846	556943	387960	352163	407192[A]	407192[A]	520703	407192[A]	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 9: Cell Means of Topcoded Earnings for Secondary Wages and Salaries (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	N/A	N/A	N/A	N/A	N/A	N/A	99999[B]	N/A	N/A	99999[B]	N/A	N/A
1989	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1990	99999[A]	N/A	N/A	N/A	N/A	N/A	99999[A]	N/A	N/A	N/A	N/A	N/A
1991	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1992	99999[B]	N/A	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1994	158174	125569[A]	N/A	125569[A]	N/A	N/A	125569[A]	N/A	N/A	N/A	N/A	N/A
1995	207148	109775[A]	109775[A]	109775[A]	N/A	N/A	109775[A]	N/A	N/A	202541	N/A	N/A
1996	64541	29777	183740	56978	35661[A]	35661[A]	77980	35661[A]	35661[A]	41501	35661[A]	N/A
1997	45749	62044[A]	62044[A]	48634	257107	62044[A]	79467	N/A	62044[A]	54580	N/A	62044[A]
1998	61345	51706	39942	48753	47529[A]	35078	88512	40732	47529[A]	39354	47529[A]	47529[A]
1999	59925	51139	52682	35583	34826	36827	50408	N/A	57830[A]	59300	57830[A]	36605
2000	50037	35625	39676	51469	67776	50770	49809	31178 [A]	31178[A]	65966	31178 [A]	236179
2001	56861	76598	39968	41433	39816	37788	55694	38587	31473	39320	37721[A]	37721[A]
2002	60672	49155	50535	43388	40556	65480	62490	37605	44090	44255	48817	37605
2003	91363	60726	49867	55255	48549[A]	57294	63216	48549[A]	49482	51519	48549[A]	N/A
2004	89986	156014	64536	67710	57296	49200	85006	50722	53836[A]	50852	53836[A]	53836[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 10: Cell Means of Topcoded Earnings for Secondary Self-Employment Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1989	125624[B]	N/A	N/A	N/A	N/A	N/A	125624[B]	N/A	N/A	N/A	N/A	N/A
1990	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	99999[B]	N/A	N/A
1991	99999[B]	N/A	N/A	N/A	N/A	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A
1992	99999[B]	N/A	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	99999[B]	N/A	N/A	N/A	N/A	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A
1994	157513[A]	N/A	N/A	157513[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1995	305001	N/A	357471[A]	357471[A]	N/A	N/A	357471[A]	N/A	N/A	357471[A]	N/A	357471[A]
1996	154533	82232[A]	N/A	64058	N/A	N/A	150320	82232[A]	N/A	82232[A]	N/A	N/A
1997	128473	152704[A]	152704[A]	152704[A]	N/A	152704[A]	152704[A]	N/A	N/A	152704[A]	152704[A]	N/A
1998	101769	N/A	104340[A]	53482	N/A	N/A	104340[A]	104340[A]	N/A	104340[A]	104340[A]	N/A
1999	123543	N/A	103546[A]	52835	N/A	N/A	103546[A]	N/A	103546[A]	131515	N/A	N/A
2000	119583	N/A	64058[A]	63258	N/A	N/A	107119	64058[A]	64058[A]	64542	N/A	64058[A]
2001	119739	59954[A]	59954[A]	59954[A]	N/A	N/A	88916	N/A	N/A	61946	N/A	59954[A]
2002	127593	108081	79681	56935	49520[A]	49520[A]	98045	N/A	49520[A]	48880	N/A	49520[A]
2003	141611	149565	149565	75881	N/A	N/A	193635	N/A	N/A	149565	N/A	N/A
2004	111644	100401[A]	91233	99461	100401[A]	N/A	154059	100401[A]	100401[A]	100401[A]	N/A	100401[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 11: Cell Means of Topcoded Earnings for Secondary Farm Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	99999[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1989	N/A	N/A	N/A	N/A	N/A	N/A	99999[B]	N/A	N/A	N/A	N/A	N/A
1990	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1991	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1992	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1994	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1996	53068	45717[A]	N/A	45717[A]	N/A	N/A	N/A	N/A	N/A	45717[A]	N/A	N/A
1997	38782[A]	N/A	N/A	38782[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1998	90173	N/A	N/A	61128	N/A	N/A	92604[A]	N/A	N/A	92604[A]	N/A	92604[A]
1999	65336	N/A	N/A	44560[A]	N/A	N/A	44560[A]	N/A	N/A	44560[A]	N/A	N/A
2000	87162	N/A	51354[A]	54785	N/A	N/A	51354[A]	N/A	N/A	62400	N/A	51354[A]
2001	133766	186210[A]	186210[A]	186210[A]	186210[A]	186210[A]	186210[A]	N/A	N/A	186210[A]	N/A	186210[A]
2002	44548	303613[A]	49415	431361	303613[A]	303613[A]	45603	N/A	303613[A]	360924	N/A	N/A
2003	65680	199339[A]	199339[A]	46838	N/A	199339[A]	107813	N/A	N/A	199339[A]	N/A	N/A
2004	67547	37498[A]	50169	40178	37498[A]	N/A	72336	N/A	301851	37498[A]	N/A	37498[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported cell mean is a combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 12: Standard Deviation of Topcoded Earnings for Wages and Salaries (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	16730	N/A	11376[A]	11376[A]	N/A	N/A	13783	N/A	N/A	N/A	N/A	N/A
1977	17185	N/A	16493[A]	15880	N/A	N/A	20139	16493[A]	N/A	N/A	N/A	N/A
1978	17100	18399[A]	18399[A]	15987	18399[A]	N/A	18651	N/A	18399[A]	18399[A]	N/A	N/A
1979	17040	18338[A]	18057	13872	18338[A]	N/A	23694	N/A	18338[A]	16376	N/A	N/A
1980	17857	8091	14581	15598	20089[A]	20089[A]	18879	N/A	11936	20089[A]	N/A	N/A
1981	17508	10027	17415	14972	9188[A]	9188[A]	16933	18115	9188[A]	15863	N/A	9188[A]
1982	10224	N/A	11276[A]	4958	N/A	N/A	8368	11276[A]	N/A	N/A	N/A	N/A
1983	10645	8083[A]	7585	8666	8083[A]	N/A	9960	N/A	8083[A]	8083[A]	N/A	N/A
1984	10613	N/A	10928[A]	10151	N/A	N/A	9440	N/A	N/A	10928[A]	N/A	10928[A]
1985	0[B]	0[B]	0[B]	0[B]	N/A	N/A	0[B]	N/A	N/A	0[B]	N/A	N/A
1986	44407	56421[A]	15531	43670	N/A	N/A	48279	N/A	N/A	56421[A]	N/A	N/A
1987	47789	35210	66902[A]	23491	36796	N/A	55464	66902	N/A	66902[A]	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 13: Standard Deviation of Topcoded Earnings for Self Employment Earnings (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	18954	N/A	9898[A]	N/A	N/A	N/A	9898[A]	N/A	9898[A]	N/A	N/A	N/A
1977	18159	20403[A]	20403[A]	20403[A]	N/A	N/A	13985	N/A	N/A	N/A	N/A	N/A
1978	19727	N/A	7812[A]	7812[A]	N/A	N/A	18808	7812[A]	N/A	7812[A]	N/A	N/A
1979	19246	6145[A]	6145[A]	6677	N/A	N/A	17235	N/A	N/A	6145[A]	N/A	N/A
1980	19372	16223[A]	11251	16223[A]	N/A	N/A	20239	N/A	16223[A]	16223[A]	N/A	N/A
1981	19846	10493[A]	18778	25342	N/A	N/A	17366	N/A	10493[A]	10493[A]	N/A	N/A
1982	10493	N/A	9087[A]	9087[A]	N/A	N/A	8913	N/A	9087[A]	9087[A]	N/A	N/A
1983	11567	9977[A]	9977[A]	10217	N/A	9977[A]	9725	N/A	N/A	N/A	N/A	N/A
1984	9913	2130[A]	2130[A]	13322	N/A	N/A	11212	N/A	N/A	2130[A]	N/A	N/A
1985	0[B]	N/A	0[B]	0[B]	N/A	N/A	0[B]	N/A	0[B]	0[B]	N/A	N/A
1986	49185	16343[A]	N/A	16343[A]	N/A	16343[A]	4600	N/A	N/A	16343[A]	N/A	16343[A]
1987	41248	N/A	66056[A]	66056[A]	N/A	N/A	40876	N/A	66056[A]	66056[A]	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 14: Standard Deviation of Topcoded Earnings for Farm Earnings (1976-1987)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1976	22387[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	22387[A]	N/A	N/A
1977	19109[A]	N/A	19109[A]	N/A	N/A	N/A	N/A	N/A	N/A	19109[A]	N/A	N/A
1978	17461	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1979	17027[A]	N/A	N/A	N/A	N/A	N/A	17027[A]	N/A	N/A	N/A	N/A	N/A
1980	15848[A]	N/A	N/A	N/A	N/A	N/A	15848[A]	N/A	N/A	N/A	N/A	N/A
1981	15544[A]	N/A	15544[A]	N/A	N/A	N/A	N/A	N/A	N/A	15544[A]	N/A	N/A
1982	7704[A]	N/A	N/A	N/A	N/A	N/A	7704[A]	N/A	N/A	N/A	N/A	N/A
1983	9163	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1984	10900[A]	N/A	N/A	N/A	N/A	N/A	10900[A]	N/A	N/A	N/A	N/A	N/A
1985	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1986	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1987	24402[A]	N/A	N/A	N/A	N/A	N/A	24402[A]	N/A	N/A	N/A	N/A	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 15: Standard Deviation of Topcoded Earnings for Primary Labor Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	60893	44656	73709	57407	51047[A]	51047[A]	40168	51047[A]	48534	42738	N/A	51047[A]
1989	56557	52129	73935	62177	70177[A]	70177[A]	70631	70177[A]	N/A	35374	70177[A]	N/A
1990	61387	56024	53451	64251	27231[A]	27231[A]	79329	27231[A]	27231[A]	12207	N/A	N/A
1991	61981	74218	33110	44146	22866	19208[A]	57583	73064	19208[A]	29491	N/A	N/A
1992	54833	39183	55207	41738	42667[A]	N/A	51506	N/A	42667[A]	36545	N/A	42667[A]
1993	55201	58042	52362	62879	12331[A]	12331[A]	53294	12331[A]	N/A	23903	N/A	12331[A]
1994	166285	179427	199169	232494	180664	59563	141874	11533[A]	115337[A]	196120	115337[A]	N/A
1995	157561	91115	162581	159107	120657	135914	110481	435396[A]	435396[A]	112271	435396[A]	N/A
1996	197084	199399	117615	141653	N/A	320652[A]	137167	320652[A]	320652[A]	393871	N/A	320652[A]
1997	236749	290168	342457	271737	272722[A]	272722[A]	213407	272722[A]	272722[A]	38308	272722[A]	N/A
1998	246514	52055	167271	224937	275049[A]	261438	228333	275049[A]	275049[A]	334697	N/A	275049[A]
1999	214453	119852	274364	267355	454248[A]	135162	307953	N/A	454248[A]	362907	454248[A]	454248[A]
2000	208901	75341	275788	193606	115846[A]	315711	188932	115846	115846[A]	115903	115846[A]	N/A
2001	252203	257132	169925	247787	239216[A]	278188	243275	N/A	239216[A]	43120	239216[A]	N/A
2002	232576	236676	231970	295423	438984	251417	184247	296601[A]	296601[A]	207485	N/A	296601[A]
2003	232238	285151	349040	317943	158086	387840	315467	N/A	N/A	186189	N/A	N/A
2004	239439	230033	257811	227899	324354	247175	200478	169098[A]	169098[A]	309739	169098[A]	N/A

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 16: Standard Deviation of Topcoded Earnings for Secondary Wages and Salaries (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	N/A	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A	0[B]	N/A	N/A
1989	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1990	0[B]	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A
1991	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1992	0[B]	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1994	72959	26262[A]	N/A	26262[A]	N/A	N/A	26262[A]	N/A	N/A	N/A	N/A	N/A
1995	83060	25576[A]	25576[A]	25576[A]	N/A	N/A	25576[A]	N/A	N/A	82629	N/A	N/A
1996	144017	1797	388997	79869	12299[A]	12299[A]	171440	12299[A]	12299[A]	15268	12299[A]	N/A
1997	46919	57662[A]	57662[A]	30638	399278	57662[A]	103526	N/A	57662[A]	45926	N/A	57662[A]
1998	99670	57214	16017	66902	27603[A]	14841	181759	15792	27603[A]	14289	27603[A]	27603[A]
1999	85348	37105	64161	14696	11196	7672	62670	N/A	45493[A]	84111	45493[A]	4734
2000	45006	9578	16224	63895	89871	20663	33643	4073[A]	4073[A]	183517	4073[A]	406968
2001	50780	90243	15559	33226	10810	14352	47357	7876	5724	19749	8198[A]	8198[A]
2002	91861	37075	57502	23261	23516	65345	81143	10280	22158	47271	18691	10280
2003	128802	47934	25949	20334	9796	14774	23638	9796[A]	11133	14762	9796[A]	N/A
2004	134647	308494	26433	36781	23428	10923	127501	11368	17919[A]	13206	17919[A]	17919[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 17: Standard Deviation of Topcoded Earnings for Secondary Self Employment Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1989	47822[A]	N/A	N/A	N/A	N/A	N/A	47822[A]	N/A	N/A	N/A	N/A	N/A
1990	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A
1991	0[B]	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A
1992	0[B]	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	0[B]	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A
1994	72219[A]	N/A	N/A	72219[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1995	308101	N/A	312669[A]	312669[A]	N/A	N/A	312669[A]	N/A	N/A	312669[A]	N/A	312669[A]
1996	247781	24028[A]	N/A	20170	N/A	N/A	110369	24028[A]	N/A	24028[A]	N/A	N/A
1997	194199	152362[A]	152362[A]	152362[A]	N/A	152362[A]	152362[A]	N/A	N/A	152362[A]	152362[A]	N/A
1998	61329	N/A	92825[A]	20340	N/A	N/A	92825[A]	92825[A]	N/A	92825[A]	92825[A]	N/A
1999	180493	N/A	26653[A]	15564	N/A	N/A	26653[A]	N/A	26653[A]	97311	N/A	N/A
2000	150800	N/A	16191[A]	18199	N/A	N/A	117147	16191[A]	16191[A]	12810	N/A	16191[A]
2001	167941	15257[A]	15257[A]	15257[A]	N/A	N/A	48195	N/A	N/A	43488	N/A	15257[A]
2002	129900	64053	103814	20367	9734[A]	9734[A]	57658	N/A	9734[A]	8139	N/A	9734[A]
2003	132875	64445	64445	4443	N/A	N/A	134913	N/A	N/A	64445	N/A	N/A
2004	70336	29081[A]	28939	43012	29081[A]	N/A	133080	29081[A]	29081[A]	29081[A]	N/A	29081[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Appendix Table 18: Standard Deviation of Topcoded Earnings for Secondary Farm Earnings (1988-2004)

	Full-Time (35+ hours per week), Full-Year (50+ weeks) Workers						*Not Full-Time, Full-Year Workers*					
	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic	Male, Not Black or Hispanic	Male, Black	Male, Hispanic	Female, Not Black or Hispanic	Female, Black	Female, Hispanic
1988	0[B]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1989	N/A	N/A	N/A	N/A	N/A	N/A	0[B]	N/A	N/A	N/A	N/A	N/A
1990	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1991	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1992	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1993	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1994	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1996	24685	12442[A]	N/A	12442[A]	N/A	N/A	N/A	N/A	N/A	12442[A]	N/A	N/A
1997	10218[A]	N/A	N/A	10218[A]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1998	57902	N/A	N/A	35884	N/A	N/A	69029[A]	N/A	N/A	69029[A]	N/A	69029[A]
1999	39783	N/A	N/A	26514[A]	N/A	N/A	26514[A]	N/A	N/A	26514[A]	N/A	N/A
2000	24565	N/A	16589[A]	10401	N/A	N/A	16589[A]	N/A	N/A	0	N/A	16589[A]
2001	252126	352636[A]	352636[A]	352636[A]	352636[A]	352636[A]	352636[A]	N/A	N/A	352636[A]	N/A	352636[A]
2002	8322	447548[A]	7909	478439	447548[A]	447548[A]	11736	N/A	447548[A]	493958	N/A	N/A
2003	56590	337119[A]	337119[A]	8864	N/A	337119[A]	86629	N/A	N/A	337119[A]	N/A	N/A
2004	67333	15364[A]	29400	3341	15364[A]	N/A	45827	N/A	420879	15364[A]	N/A	15364[A]

[A] - Indicates that there are not enough observations to report a cell mean for this population group in this year. The reported standard deviation is the standard deviation for the combined cell mean with the other footnoted population groups in the same year.
[B] - Indicates that the internal censoring point is identical to the public cell mean so no additional information can be obtained from the internal data.
N/A – Indicates that no individual with these demographic characteristics were topcoded in this year from the specified income source

Source: Author's calculations using internal March CPS data

Endnotes

[1] Labor earnings include income from: primary earnings, wages and salaries, self-employment, and farm-earnings sources. This is a slightly broader definition of earnings than income from primary earnings and wages and salaries that is most commonly used in the literature.

[2] Although researchers have explored hourly, weekly, and annual wage inequality using March CPS data, all of these measures of wages are based on respondents reported annual income from wages and salaries. Wage earnings are then rescaled to a weekly or hourly basis by dividing annual income from wages and salaries by the weeks or hours worked over the course of a year.

[3] Each year the Census Bureau uses the internal March CPS data to produce its official statistics on the average income and poverty rates (U.S. Census Bureau, 2007b). These official values are based on the internal March CPS data that is not available, except under certain conditions, to researchers outside of the Census Bureau.

[4] Each CPS survey captures income from the previous year. In this paper all references are to the survey year, so when we discuss the year 1976, this refers to income from various sources that members of the household received in 1975 reported on the March 1976 Current Population Survey.

[5] In the analysis of the log variance of earnings, we use the cell means and variances for all individuals from each source of topcoded labor earnings without dividing the sample by demographic information. This provides more stable parameter estimates, since the small number of topcoded individuals in some demographic groups would lead to large fluctuations in these estimates.

[6] Because of the complexity of these equations, an analytical solution does not exist to solve the simultaneous equations. Instead, the solution to the moment equations were calculated using a constrained minimization problem, setting the first moment equation as a constraint and finding α and θ that minimize the square of the second moment equation with the additional constraint that both α and θ are in the positive domain. To ensure that the solution found through the constrained minimization problem solves both moment conditions and is not simply the best available solution given the constraints, a similar minimization was repeated using the second moment as the constraint and minimizing the square of the first moment equation. In all years, we found identical solutions of α and θ when performing the second minimization problem.

[7] Since the Pareto distribution is a special case of the Stoppa distribution, it is easy to test whether the Pareto assumption that the θ parameter equals 1. Unfortunately, due to the relatively small number of topcoded individuals from labor

earnings sources in some years, for confidentiality reasons on the internal data we are unable to release the variance-covariance matrices that would be required to produce standard errors for our cell means and variances in order to test whether the Pareto assumption is satisfied in specific years.

[8] In creating the Public-Multiple series, we followed an approach similar to that used by Lemieux (2006). For self-employment and farm earnings in all years, we assign topcoded individuals earnings from that source equal to 1.4 times the topcode threshold. For wages, prior to 1988, topcoded individuals for each source of income are assigned 1.4 times the topcode threshold. After the 1988 Census redesign, we follow Lemieux's approach and assume that all primary labor earnings are from wages and salaries and thus combines the primary labor earnings and secondary earnings from wages and salaries. From 1988 to 1995 individuals with combined wages greater than 99,999 from the two sources are considered topcoded and assigned a value of 1.4 times the topcode threshold of 99,999. Starting in 1996, individuals with combined wages greater than 150,000 from the two sources, which is the topcode threshold for primary labor earnings, are considered topcoded and assigned a value of 1.4 times the topcode threshold of 150,000.

[9] Note that while imputing topcoded values using the cell means and variances allows us to better match measures of earnings and income inequality using the public-use CPS to results found when using the internal CPS data, these methods are limited in detecting the true levels of US inequality because the internal CPS data is also subject to a degree of censoring. The internal censoring points are generally much higher and more stable over time than the public-use topcodes, but this censoring means that the very top values will still not be captured. Additionally, a substantial increase in internal censoring points in 1994 leads to a trend break in that year for any series based on the internal CPS. Despite these limitations, since this is the same data used to produce official Census Bureau statistics, these remaining censorship problems for users of the variance and cell mean data in conjunction with the public-use March CPS data are no worse than those in the data used by the Census Bureau to produce its official statistics with the internal March CPS.