On the Reliability of Software Piracy Statistics

I.P.L. Png*

Revised, September 2008

Abstract

Despite tremendous debate and policy interest in software piracy, statistics compiled by the Business Software Alliance (BSA) have generally been accepted at face value by policy makers and scholars. However, the accuracy of BSA statistics has not been independently verified.

Based on a review of the BSA methodology and empirical analysis, I conclude the following. A change in the BSA consultant and methodology around 2002-03 had systematic effects on published piracy rates. First, the trend rate of decrease of piracy rates fell from 2.0% points per year to 1.1% points per year. Second, piracy rates apparently became more sensitive to changes in income. Third, piracy rates depended on projections of software usage based on per capita incomes in the respective countries.

* NUS Business School, National University of Singapore, 1 Business Link, Singapore 117592, Tel: +65 6516-6807; http://www.comp.nus.edu.sg/~ipng/. I thank Hui Kai-lung, Barrie Nault, V. Sambamurthy, Qiu-hong Wang, and participants at seminars at the Research Institute for Economy, Trade, and Industry, Japan, for helpful discussions and advice.

1. Introduction

Major copyright-protected industries -- business software, movie, music, electronic games, books, broadcasting -- publish statistics of piracy against their respective industries. These statistics play an important role in public policy. They are systematically cited by industry to argue for more government enforcement and increased penalties against piracy.

Piracy statistics play a central role in U.S. international trade policy. Special 301 of the U.S. Trade Act of 1974 requires the U.S. Trade Representative to report annually on countries that do not provide adequate and effective protection of intellectual property (IP) rights, or deny fair and equitable market access to U.S. exporters of IP-protected items. Countries whose laws, policies, or practices are deemed to adversely affect U.S. producers or products may be subject to investigation, trade sanctions, or other penalties. Piracy statistics reported by the International Intellectual Property Alliance (IIPA) are the key evidence in Special 301 reviews (see, for instance, IIPA (2007)).

Piracy statistics are regularly cited in free trade negotiations between the United States and other countries. For instance, Table 7-1 of the U.S. International Trade Commission report on the economic effects of the U.S.-Chile free trade agreement quotes IIPA statistics of piracy in Chile for 2002: movies -- 40%, recorded music -- 35%, business software -- 51%, and electronic games -- 78%.[1]

[1] U.S. International Trade Commission, U.S.-Chile Free Trade Agreement: Potential Economywide and Selected Sectoral Effects, Investigation No. TA-2104-5, USITC Publication 3605, June 2003.

The U.S. Trade Representative cites industry piracy statistics in its annual reports on foreign countries. For instance, the 2007 report on Singapore states: "According to industry estimates, Singapore's piracy rate averaged 5 percent for music and 12 percent for movies. Business software losses were estimated at nearly $86 million in 2005".[2]

The public-policy implications of piracy have motivated numerous academic studies into the causes of piracy. These have focused on business software piracy and investigated how it differs across countries (e.g., Gopal and Sanders 1998 and 2000; Marron and Steel 2000; Hufsted 2000; Depken and Simmons 2004; van Kranenburg and Hogenbirk 2005; Rodriguez 2005; Fischer and Rodriguez 2005; Rodriguez 2006; Chellappa et al. 2006; Robertson et al. 2007) and within the United States (Bezmen and Depken 2006).[3]

The widespread use of piracy statistics published by the Business Software Alliance (BSA) in public policy and academic research prompts the question: How accurate and reliable are these statistics, and in particular, are they biased in any way?

Here, I review the methodology, coverage, and implementation of the BSA software piracy statistics and previous studies of the causes of software piracy, and conclude that a change in the BSA consultant and methodology around 2002-03 had systematic effects on published piracy rates. First, the trend rate of decrease of piracy

[2] U.S. Trade Representative, 2007 report on Singapore, http://www.ustr.gov/assets/Document_Library/Reports_Publications/2007/2007_NTE_Report/asset_upload_file129_10979.pdf [Accessed, November 10, 2007].

[3] Other academic studies have used piracy to explain various business and marketing strategies and public policy issues – including sales of recorded music (Hui and Png 2003), enforcement by software publishers (Gu and Mahajan 2005), and the incentive to produce movies (Png and Wang 2007).

rates fell from 2.0% points per year to 1.1% points per year. Second, piracy rates apparently became more sensitive to changes in income. Third, piracy rates depended on projections of software usage based on per capita incomes in the respective countries.

2. Methodology

Copyright law governs the expression contained in, *inter alia*, books, music, movies, software, electronic games, databases, and designs. Industry associations for producers of business software, music, electronic games, and movies publish international statistics of piracy. The statistics differ in the scope of coverage, frequency of publication, methodology, and sources of primary information.

Table 1 reports various aspects of the piracy statistics published by the four industry associations.[4] The statistics published by the Business Software Alliance (BSA) provide the widest geographical coverage, are published annually, and, appear to be the most transparent in terms of methodology, data sources, and implementation. The methodology is published in some detail, uses both internal and external data, and, since 2003, has been implemented by a well-reputed consultant, viz., International Data Corporation (IDC) (BSA 2004).

--- Table 1. Industry piracy statistics ---

For 2002 and earlier years, the business software piracy statistics were produced by the International Planning and Research Corporation (IPRC) (BSA 2003). IPRC

[4] This Table relies, in part, on Hui and Png (2005).

focused on three groups of business PC software – general productivity applications, professional applications, and utilities.

The IPRC estimated piracy using an indirect methodology. For each country, the quantity of pirated software was estimated as being the difference between the quantity installed and the quantity legitimately acquired. In turn, the quantity installed was estimated as the number of computers in use multiplied by corresponding norms for the "software load" in four customer segments -- new and existing residential computers, and new and existing business computers.

Software load is the quantity of software installed per computer. The norms for software load for the four segments were based on U.S. market research (BSA 2003, page 11). However, the IPRC did not explain whether, and if so, how it adjusted the U.S. norms to compute the software load in other countries.

The IPRC directly estimated the numbers of computers in use "for the major countries … from proprietary and confidential data supplied by BSA member companies", while "[t]he "rest of region" data was used to develop piracy estimates outside of the major markets" (BSA 2003, pp. 11-12). The IPRC did not specify the "major" markets or method by which it developed the "rest of region" data.

In 2003, BSA engaged a new consultant, IDC, and the methodology underlying the business software piracy statistics was refined. IDC expanded the scope of measurement to cover all PC software, including operating systems, systems software such as databases and security packages, and general and specific applications software.

The IDC applied the same basic methodology, estimating the quantity of pirated software indirectly as the difference between the quantity installed and the quantity legitimately acquired. In turn, the quantity installed was estimated as the number of computers in use multiplied by norms for the software load.

By contrast with the IPRC which applied U.S. norms for the software loads, the IDC calculated the software loads per new computer and per existing computer from surveys of consumers and business users in 15 countries -- Bolivia, Brazil, Chile, China, Colombia, Costa Rica, Dominican Republic, Guatemala, Kuwait, Malaysia, Mexico, Romania, Spain, Taiwan and the United States (BSA 2006, page 17).[5]

IDC explained that the "results of these surveys were used to populate IDC's input models for the other countries. For 2004 and 2005, IDC updated these models based on additional local country research, interviews with users and the channels, and spot inventories" (BSA 2004, page 11). However, IDC did not explain whether, and if so, how it adjusted the norms from the 15 countries to compute the software load in the other countries.[6]

As for the numbers of computers, IDC collected information on PC shipments for "more than 75 countries", while for the "additional 25-plus countries and markets, the data were either collected in-country or modeled regionally based on IDC's rest-of-region estimates" (BSA 2004, page 10). The IDC did not elaborate on how it "modeled" the

[5] IDC did not explain the selection of the 15 countries for calculating the software loads. In terms of geography, Europe and Asia seem under-represented, while Central and South America seem over-represented. In terms of economic development, advanced countries seem under-represented and less developed countries seem over-represented.

[6] Subsequently, the IDC surveyed 21 countries in 2006, and 22 countries in 2007 (BSA 2008). However, the IDC did not disclose the identity of the surveyed countries.

number of computers in the other 25-plus countries. Nevertheless,,the IDC was more transparent with regard to methodology (5 pages) than the previous consultant, IPRC (just 2 pages).[7]

From the viewpoint of policy and research, the key questions are: (i) whether the change in consultant from IPRC to IDC, and the revision of the estimation methodology had any systematic impact on piracy statistics; and (ii) whether the methods used by BSA consultants to project the software load and number of computers from countries in which they compiled these data to other countries had any systematic impact on piracy statistics.

To address these questions, I compiled national piracy rates for 81 to 103 countries over the eleven-year period 1997-2007 from BSA publications (the panel began with 81 countries in 1997 and ended with 103 countries in 2007). The period of study included seven years before and five years after the change in consultant and methodology. Accordingly, it would provide good coverage of any impact of the change. For brevity, I refer to the years, 1997-2002, before the change as "pre-change" and the years, 2003-07, after the change as "post-change".

3. Trend

Table 2 reports various tests of the impact of the change in consultant and methodology on national piracy rates. Specifications (i) and (ii) were simple ordinary least-squares

[7] In particular, the IPRC was silent on whether they actually measured the numbers of each segment of computers in every country, or somehow projected these numbers as the IDC did.

(OLS) regressions of the piracy rate on year indicators and pre- and post-change year trends respectively.[8] The results from both specifications suggest that piracy rates were subject to a significant downward trend.

Referring to specification (i), the year indicators show a clear decreasing trend in the average piracy rate from 1999-2002. The results from specification (ii) are even clearer: prior to the change in consultant and methodology, the average piracy rate fell by 2.0% points annually, while, after the change, the average rate fell at the slower rate of 0.7% points annually. The difference between the pre-change and post-change trends was statistically significant ($F(1, 989) = 15.30$, $p < 0.0001$).

--- Table 2. BSA change of consultant and methodology ---

The OLS regression pooled all countries, regardless of economic, institutional, or cultural differences, into a single estimate. Obviously, it would be more appropriate to account for any systematic national differences. Specifications (iii) and (iv) replicated the analysis, using country fixed effects. The fixed effects would account for any systematic national differences which did not vary over time. The results were even stronger than with the OLS estimates.

Referring to specification (iii), the year indicators show a clear decreasing trend from 1998-2002, and then an upward shift by about 1.7% points between 2002-03. Referring to specification (iv), the trend rate of decrease of piracy rates was significantly

[8] The data exhibited serial correlation within panels ($F(1, 102) = 142.78$, $p < 0.0001$, using the test proposed by Wooldridge (2002), page 282, as implemented by Drukker (2003)). Accordingly, all standard errors for fixed-effects estimates in Tables 2, 5, and 6 were calculated by country-level cluster.

higher before the change in consultant and methodology (–2.0% points per year) than after the change (–1.1% points per year). The difference between the pre-change and post-change trends was statistically significant (F (1, 102) = 60.18, $p < 0.0001$).

One possible reason why the downward trend of piracy rates decelerated around 2002-03 was the expansion of BSA coverage to include countries with higher piracy rates. Specifically, in 1997, the BSA piracy statistics covered 81 countries, while in 2007, they covered 103 countries. The coverage was expanded to include countries with relatively high piracy, such as Albania (78%), Kazakhstan (79%), and Zambia (82%).[9]

To avoid any bias due to the expanded coverage, specifications (v) and (vi) limited the fixed effects estimates to those countries covered throughout the period, 1997-2007. The results from the balanced sample were similar. Figure 1 illustrates the evolution of the average piracy rate, by year. Referring to specification (vi), the pre-change trend rate of decrease of piracy rates (–2.1% points per year) was significantly higher than the pre-trend change (–1.1% points per year) (F(1, 80) = 58.55, $p < 0.0001$).

--- Figure 1. Average piracy rate ---

Could the change in the trend decrease of piracy rates be due to changes in some factor other than the BSA's change in consultant and methodology? In the sections below, I present various robustness checks to rule out this possibility. Specifically, I included other covariates which have been shown to influence software piracy – income per capita, national culture, and institutions. I found the same conclusion: the trend rate

[9] Numbers in parentheses are the respective piracy rates for 2007.

of decrease of piracy rates was significantly higher before the change in consultant and methodology than after the change.

4. Software Load and Numbers of Computers

Formally, the methodology applied by BSA consultants, IPRC and IDC, was to estimate the quantity pirated in country *i* during year *t* as the difference between usage of software, U_{it}, and the quantity of software legitimately acquired, S_{it},

$$P_{it} = U_{it} - S_{it}. \quad (1)$$

In turn, the usage was computed as

$$U_{it} = \lambda_{bni} B_{nit} + \lambda_{bei} B_{eit} + \lambda_{hni} H_{nit} + \lambda_{hei} H_{eit}, \quad (2)$$

where λ_{bni}, λ_{bei}, λ_{hni}, and λ_{hei} were norms for software load on new and existing computers for business and home users in country *i* respectively, and B_{nit}, B_{eit}, H_{nit}, H_{eit}, were the numbers of new and existing computers in use among business and home users in country *i* respectively.

The piracy *rate* in country *i* for year *t* was then calculated as the ratio of the pirated quantity to usage,

$$r_{it} = \frac{P_{it}}{U_{it}} = \frac{U_{it} - S_{it}}{U_{it}} = 1 - \frac{S_{it}}{U_{it}}. \quad (3)$$

As noted above, neither IPRC nor IDC revealed whether, and if so, how they adjusted the norms for software load from the sampled countries to compute the norms in other countries. Further, IDC did not reveal how it projected the number of computers in

the sampled countries to other countries. IPRC did not even disclose whether it used any projection.

Intriguingly, Marron and Steel (2000) remarked: "the trade groups' estimation procedure involves significant assumptions. This raises the question of whether our empirical analysis might uncover artifacts of the estimation procedure rather than true relationships among the variables. For example, *did the analysts assume that high-income countries have lower piracy rates*?" (footnote 8, emphasis added). However, this crucial question was not formally addressed.[10]

With instrumental variable methods, it is straightforward to address this issue. Re-arranging equation (3), the rate of *legitimate* consumption is

$$c_{it} = 1 - r_{it} = \frac{S_{it}}{U_{it}}. \qquad (4)$$

In logarithms, this becomes

$$\ln c_{it} = \ln S_{it} - \ln U_{it}. \qquad (5)$$

Formally, the issue is whether national income per capita is exogenous in a regression of legitimate consumption rates (or equivalently, piracy rates) on national income per capita. With a slight abuse of notation, let $\ln Y_{it}$ represent a vector of the logarithm of variables relating to income per capita.

Now, suppose that the BSA consultant projected software usage according to

$$\ln U_{it} = \beta \ln Y_{it}, \qquad (6)$$

[10] Marron and Steel (2000) dismissed the concern "[a]fter speaking with the consultant who prepared the trade groups' figures" (footnote 8).

where $\beta > 0$ is a vector of coefficients. Substituting from (6) in (5),

$$\ln c_{it} = \ln S_{it} - \beta \ln Y_{it}.$$

So, a regression of $\ln c_{it}$ on $\ln Y_{it}$ would fit perfectly and yield $-\beta$ as the coefficient of $\ln Y_{it}$.

Suppose, however, that the BSA consultant projected usage according to

$$U_{it} = U(Y_{it}), \qquad (7)$$

where $U_{it} = U(.)$ is not linear or a power function. Then, substituting from (7) in (5),

$$\ln c_{it} = \ln S_{it} - \ln U(Y_{it}), \qquad (8)$$

and, hence, in any regression of $\ln c_{it}$ on $\ln Y_{it}$, the error term would be correlated with the dependent variable. Equivalently, income would not be exogenous. The essential reason is that software usage is projected on the basis of income and the projection does not conform to the specification, (6), assumed in the regression.

A Hausman test (Wooldridge (2006), pp. 532-533) of the endogeneity of income in equation (10) would proceed as follows. Perform first stage regressions for every income variable – regressing every income variable on instruments for the income variables and exogenous variables that affect piracy. In the second stage, regress the rate of legitimate consumption on the income variables, exogenous variables that affect piracy, and the residuals from the first-stage regressions. If any of the first-stage residuals turns out to be significant in the second-stage regression, then the corresponding first-stage regressor is not exogenous. The implication would be that, indeed, legitimate consumption and hence piracy rates were based on income.

-- Table 3. Descriptive statistics --

-- Table 4. Correlations --

Tables 3 and 4 present descriptive statistics of the data and pairwise correlations respectively. The key issue in any instrumental variables analysis is the choice of instruments. I identified per capita residential electricity consumption and exports as possible instruments. Residential electricity consumption would depend on income, but, intuitively, there does not seem to be any direct relation between electricity consumption and legitimate consumption of software, except to the trivial extent that software is used on computers and computers consume electricity. Similarly, income would increase with exports (indeed, "export-led" growth is a major theme in economic development), but, intuitively, there does not seem to be any direct relation between exports and legitimate consumption of software.

Referring to Table 4, both residential electricity consumption and exports were relatively more correlated with income than with piracy.[11] Since legitimate consumption is the complement of piracy, this would also imply that both residential electricity consumption and exports were relatively more correlated with income than with legitimate consumption.

Table 5 reports tests of endogeneity, using post-change residential electricity consumption per capita as the instrument for post-change income per capita. (In Tables 5 and 6, except as otherwise noted, all variables except year trends were specified in

[11] Referring to Table 4, the correlation between electricity consumption and piracy was -0.589. This is probably an indirect correlation due to the separate correlations between electricity consumption and income (0.765) and between income and piracy (-0.815).

logarithms (Wooldridge (2006), pp. 197-200), and all regressions included country fixed effects, and to be conservative, were limited to the balanced panel.)

--- Table 5. Endogeneity of income ---

Referring to Table 5, columns (ii) and (iii), in the first-stage regression of post-change income, the coefficient of post-change electricity consumption was significant, while in the second-stage regression, the coefficient of the residuals from the first-stage regression was significant. This suggests that, over the years 2003-07, following the change in consultant and methodology, income was endogenous.[12] The results were similar using post-change exports per capita as the instrument, which, for brevity, I do not report here.

Table 5, column (iv), reports the 2SLS (2-stage least squares) estimates. By comparison with the OLS estimate in column (i), the major differences were that the coefficient of post-change income and the post-change trend were about double in magnitude. These differences show that the BSA consultant's projections of software usage affected the post-change trend as well as the coefficient of post-change income.

Referring to Table 5, column (iv), observe that the coefficient of post-change income was positive and significant (0.095 (±0.026)). This result is robust – the coefficient of post-change income was also positive and significant in the OLS regression and in the 2SLS regression using exports as an instrument. Apparently, not only did the change in consultant and methodology affect the trend in piracy rates, but it also caused

[12] In the similar Hausman test (unreported) of both income and post-change income, only the residuals of post-change income were significant in the second stage. Accordingly, I inferred that only post-change income was endogenous and that income was not endogenous.

legitimate software consumption to become more sensitive to income per capita. There seems to be no obvious reason for this except the assumptions underlying the BSA consultant's projection of software usage.

The significance of post-change income also helps to confirm that the endogeneity of income was due to the consultant's projection of software usage. An alternative explanation – that both income and usage were related to some omitted factor (e.g., computer literacy or education) – is implausible as there is no obvious reason for the relation between usage and the hypothetical omitted factor to have changed around 2002-03.[13] [14]

I infer that, following the change in consultant and methodology in 2003, income was *not exogenous* to software piracy. This finding was consistent with the norms for software load and the numbers of computers being projected on the basis of income per capita. Indeed, IDC disclosed that it based the norms for software load on a study of 15 countries and that it "modeled" the numbers of computers in some countries (BSA 2004, pp. 10-11). [15]

---

[13] In robustness checks below, I included these other variables and found that they did not affect the finding that post-change income was endogenous.

[14] Referring to Table 5, column (iv), the pre-change trend in *legitimate consumption* was positive, while the post-trend change was negative. Accordingly, before the change in consultant and methodology, the average *piracy* rate systematically decreased over time, while, after the change, the average piracy rate systematically increased over time. This is even stronger than the results presented in Table 2.

[15] Husted (2000) considered that BSA statistics to be reliable for cross-country study as there was "no indication that BSA has an interest in demonstrating lower rates of piracy in certain areas rather than in others" (pp. 207-208). He skirted the issue of whether there was any systematic bias in how BSA consultants projected software loads and numbers of computers from sampled countries to other countries. Moreover, he apparently overlooked the use of BSA statistics in U.S. government reports and investigations of intellectual property violations to target specific countries under Special 301.

By contrast, I did not find any evidence that income was endogenous to published rates of software piracy between 1997-2002. Apparently, the responsible consultant, IPRC, simply applied U.S. software norms to other countries, without any adjustment (BSA 2003, page 11).[16]

Without access to the BSA's estimates and independent information on software usage, it is not possible to comment on the *accuracy* of the BSA's projections of software usage and the consequent estimates of piracy rates. All that can be said is that the BSA's consultant did project software usage on the basis of income from 2003 onward. Perhaps the projection was accurate, or perhaps it was not.

5. Robustness

In previous studies of the determinants of business software piracy (Gopal and Sanders 1998 and 2000; Husted 2000; Marron and Steel 2000; Shin et al. 2004; Depken and Simmons 2004; van Kranenburg and Hogenbirk 2005; Rodriguez 2006; Fischer and Rodriguez 2005; Bagchi et al. 2006; Chellappa et al. 2006; Robertson et al. 2007), three factors stood out as being robust to alternative specifications, geographical coverage, and time periods.[17] They were national income per capita, individualism (a dimension of national culture), and law and institutions.

Could the variation in one or more of these factors possibly explain the changes in the pattern of piracy rates and their relation with income around 2002-03 or the

[16] IPRC did not disclose its methodology with regard to the numbers of computers in the various countries.

[17] Of the multiple studies, only Fischer and Rodriguez (2005) included country fixed effects. All others applied cross-section analysis without considering variation over time.

endogeneity of income? In the following discussion, I consider the impact of including these other factors relative to the 2SLS estimate in Table 5, column (iv), which is treated as a benchmark.

The results reported in Table 5 suggest that changes in income could not account for the changes in the pattern of piracy rates around 2002-03. Indeed, the empirical results suggest the opposite – that the relationship between piracy rates and income itself changed around 2002-03.

The sociologist, Geert Hofstede (1983, 2001) famously developed four indexes of national culture from surveys of IBM employees over the period 1967–1973 in 40 countries. The indexes were later extended to other countries. One index measured individualism (and its complement, collectivism). Hofstede's (1983, 2001) indexes are static. Accordingly, the individualism index could not be used to account for changes in piracy rates or their relation with income over time. However, could differences in individualism account for the endogeneity of income?

Table 6, column (a), reports a 2SLS regression of legitimate consumption including individualism as an additional covariate.[18] To complete each country's panel for individualism, the same value was replicated for every year, hence, fixed effects could not be used. In the estimates, legitimate consumption was significant and increasing in individualism. The coefficient of post-change income was positive, about one-third of the magnitude as in the benchmark, but was not precisely estimated. This imprecision might be due to the high correlation between income and individualism (Table 2).

[18] All 2SLS regressions presented in Table 6 used post-change residential electricity consumption per capita as the instrument for post-change income per capita.

Moreover, with individualism being constant over time, I could not use fixed effects, hence, the estimates were driven by across-country variation as well as within-country variation.[19] For these reasons, the estimate including individualism should be treated with caution.

--- Table 6. Robustness ---

With regard to law and institutions, the obvious measure is the World Bank's rule of law index, which was also used by Fischer and Rodriguez (2005) and Chellappa et al. (2006). This index is compiled from multiple primary sources, and is a perceptual measure of "the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence" (Kauffman et al. 2007, page 4). The index was compiled biennially from 1996-2000, and annually from 2002 onward.

Table 6, column (b), reports a 2SLS regression of legitimate consumption including the rule of law index. The results were qualitatively similar to the benchmark estimate: the coefficient of post-change income was positive and significant, and the average consumption was subject to an increasing trend before the change in consultant and methodology but a decreasing trend afterward. The rule of law itself was marginally significant.[20]

[19] Fixed-effects estimates are essentially estimates relative to the mean value of the dependent variable for each country, and hence, driven by within-country variation.

[20] One possible reason why the coefficient of income was lower in this estimate than in the benchmark is that improvements in the rule of law might raise both income and legitimate consumption of software. Hence, in the benchmark estimate, with the rule of law excluded, the income variable would absorb part of the effect of the rule of law.

In the previous section 4, I mooted the possibility that post-change income appeared to be endogenous in regressions of legitimate consumption because usage was determined by an omitted factor, such as computer literacy or education, which was correlated with income. Table 6, columns (c)-(d), reports 2SLS regressions of legitimate consumption including the percentage of personal computer users among the population and the literacy rate respectively. The results were very close to benchmark: the coefficient of post-change income was positive and significant, and average consumption was subject to an increasing trend before the change in consultant and methodology but a decreasing trend afterward. Neither computer usage nor the literacy rate was significant.[21]

All 2SLS regressions presented in Table 6 used post-change residential electricity consumption per capita as the instrument for post-change income per capita. I obtained very similar results (for brevity, unreported) with exports per capita as the instrument.

6. Conclusions

U.S. government pronouncements and actions as well as many academic studies have taken BSA software piracy statistics at face value. Based on a review of the BSA methodology and empirical analysis, I conclude that a change in the BSA consultant and methodology around 2002-03 had systematic effects on published piracy rates.

[21] In another unreported test, I obtained similar results using the percentage of Internet users among the population.

- The trend rate of decrease of piracy rates fell from 2.0% points per year to 1.1% points per year, so, raising piracy rates from the levels that would have been implied had they followed the trend before the change.
- Piracy rates apparently became more sensitive to changes in income.
- Piracy rates depended on projections of software usage based on per capita incomes in the respective countries.

These results were robust to the sample of countries, choice of instruments, and inclusion of alternative explanatory variables.

The key direction for future research is to gain access to the BSA methodologies and data so as to better understand the biases in their statistics, and so that future policy and research can be appropriately calibrated. Meanwhile, the central implication of my analysis is that BSA statistics should be used with great caution. Any government pronouncement or action, and study of software piracy should take account of changes in the BSA consultant and methodology, and the projections underlying the published piracy rates.

References

Bagchi, Kallol, Peeter Kirs, Robert Cerveny, "Global software piracy: can economic factors alone explain the trend? *Communications of the ACM*, Vol. 49 No. 6, June 2006, 70-76.

Bezmen, Trisha L., and Craig A. Depken II, "Influences on software piracy: Evidence from the various United States", *Economics Letters*, Vol. 90, No. 3, March 2006, 356-361.

Business Software Alliance, *First Annual BSA and IDC Global Software Piracy Study*, 2004.

Business Software Alliance, *Third Annual BSA and IDC Global Software Piracy Study*, 2006.

Business Software Alliance, *Fifth Annual BSA and IDC Global Software Piracy Study*, 2008.

Business Software Alliance, *Eighth Annual BSA Global Software Piracy Study: Trends in Software Piracy, 1994-2002*, June 2003.

Chellappa, Ramnath K., Yuanyuan Chen, and Sriram Venkataraman, "An Empirical Examination of Global Software Piracy: Pricing and Policy Implications", Goizueta Business School, Emory University, Atlanta GA 30322, 2006.

Depken II, Craig A., and Lee C. Simmons, "Social construct and the propensity for software piracy", *Applied Economics Letters*, Vol. 11 No. 2, February 2004, 97–100.

Drukker, David M., "Testing for serial correlation in linear panel-data models", *Stata Journal*, Vol. 3 No. 2, 2003, 168–177.

Fischer, Justina A.V., and Antonio Rodríguez Andrés, "Is Software Piracy a Middle Class Crime? Investigating the Inequality-Piracy Channel", Discussion paper no. 2005-18, Department of Economics, University of St. Gallen, August 2005.

Gopal, Ram D., and G. Lawrence Sanders, "Global software piracy: you can't get blood out of a turnip", *Communications of the ACM*, Volume 43, Issue 9, September 2000, 82 – 89.

Gopal, Ram D. and G. Lawrence Sanders, "International Software Piracy: Analysis of Key Issues and Impacts", *Information Systems Research*, Vol. 9 No. 4, December 1998, 380-397.

Gu, Bin, and Vijay Mahajan, "How Much Anti-Piracy Effort is Too Much? – A Study of the Global Software Industry", Red McCombs School of Business, University of Texas at Austin, 2005.

Hofstede, Geert, "Dimensions of national cultures in fifty countries and three regions", in J. B. Deregowski, S. Dziurawiec and R. C. Annis, eds., *Expications in Cross-Cultural Psychology*, Lisse, Netherlands: Swets and Zeitlinger, 1983, 333–355.

Hofstede, Geert, *Culture's Consequences, Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2nd edition, Thousand Oaks, CA: Sage Publications, 2001.

Hui, Kai-Lung, and Ivan Png, "Piracy and the Legitimate Demand for Recorded Music", *Contributions to Economic Analysis & Policy*, Vol. 2 No. 1, 2003, Article 11.

Hui, Kai-Lung, and Ivan Png, "Measuring Piracy of IP-Protected Goods: A Methodology", World Intellectual Property Organization: Working Group of Experts on Developing a Methodology for Measuring Copyright Piracy, November 1-4, 2005, http://www.econaa.com/resources/methodology_piracy_rate_sector.pdf [Accessed, July 27, 2006].

Husted, Bryan W., "The Impact of National Culture on Software Piracy", *Journal of Business Ethics*, Vol. 26 No. 3, August 2000, 197-211.

International Intellectual Property Alliance (IIPA), Press Release: The Copyright Industries in the International Intellectual Property Alliance (IIPA) Submit to USTR their 2007 Report on Piracy in 60 Countries/Territories, Washington, DC, February 17, 2007.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Governance Matters IV: Aggregate and Individual Governance Indicators 1996–2006", Policy Research Working Paper 4280, World Bank, July 2007.

Marron, Donald B., and David G. Steel, "Which Countries Protect Intellectual Property? The Case of Software Piracy", *Economic Inquiry*, Vol. 38 No. 2, April 2000, 159-174.

Png, I.P.L., and Qiuhong Wang, "Copyright Duration and the Supply of Creative Work: Evidence from the Movies", Dept of Information Systems, National University of Singapore, revised, November 2007.

Robertson, Christopher J., K. Matthew Gilley, Victoria Crittenden, and William F. Crittenden, "An analysis of the predictors of software piracy within Latin America", *Journal of Business Research*, forthcoming 2007.

Rodríguez, Andrés Antonio, "The relationship between copyright software protection and piracy: Evidence from Europe", *European Journal of Law and Economics*, Volume 21, Number 1, January 2006, 29-52.

Schaffer, Mark E., "xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models", http://ideas.repec.org/c/boc/bocode/s456501.html, 2007.

Shin, Seung Kyoon, Ram D. Gopal, G. Lawrence Sanders, and Andrew B. Whinston, "Global software piracy revisited", *Communications of the ACM*, Vol. 47 No. 1, January 2004, 103-107.

van Kranenburg, Hans, and Annelies Hogenbirk, "Multimedia, Entertainment, and Business Software Copyright Piracy: A Cross-National Study", *Journal of Media Economics*, Vol. 18, No. 2, April 2005, 109-129.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press, 2002.

Wooldridge, Jeffrey M., *Introductory Econometrics: A Modern Approach*, 3d edition, Mason, OH: Thomson Southwestern, 2006.