REVIEW OF METHODOLOGY

NATIONAL ALLIANCE FOR JOBS AND INNOVATION

NATIONAL ASSOCIATION OF Manufacturers

Appendix A

This Appendix provides detail on our methodology to estimate the impact of global software piracy on U.S. manufacturers' employment and earnings. To estimate the cost to the U.S. manufacturing sector of global piracy in year t and for economic impact variable i (i = *Employment, Revenue, GDP*), we multiply five quantities as set forth in the equation:

$$Cost\ of\ Piracy_{it} = Piracy\ rate_t * Software\ cost_t * Supply\ chain\ adjustment * Price\ elasticity_i * Impact\ variable_{it}$$

This calculation can be thought of as having three steps. The first is to measure the impact of piracy rates on prices. This is determined in the top row by multiplying the piracy rate times the software cost share, with an adjustment for the supply chain. This cost advantage is then scaled into its relative impact for an economic impact variable using the price elasticity. Finally, the relative impact is scaled by the total size of the impact variable. We describe below the calculation of each component.

Global piracy rate

The first step in estimating the size of the cost advantage resulting from global software piracy is to estimate the average rate of software piracy for global manufacturing. We obtain data on piracy rates by country in 2003 (the earliest year available) from the Business Software Alliance. We do not have data that allow us to observe variation in piracy across manufacturing industries; therefore, we estimate a single global weighted average piracy rate for each year. We obtain our weights from UN Comtrade data on global exports, which reports the dollar value of each country's global exports by year. Weighting each country's piracy rate by its global export share at five-year intervals, we arrive at the global weighted average piracy rate (see Report Figure 2).

Estimating the cost of global piracy requires modelling a scenario in which piracy is limited or reduced. As reducing piracy rates to zero is likely not feasible, we instead estimate the cost of global piracy over and above the U.S. rate of compliance. BSA reports a rate of 22% for the U.S.; we therefore subtract 22% from the annual weighted average global rate estimated above to arrive at the residual rate of global piracy that represents a cost advantage over U.S. firms.

Software expenditure by manufacturers

We next estimate manufacturers' expenditure on software as a percentage of total cost. The most detailed data available are kept by International Data Corporation (IDC), who estimate and report annual expenditure on packaged software by the U.S. manufacturing sector as a whole. We divide software expenditure by the annual total value of shipments (as reported by the Census) to arrive at the required percentage value. As described in the report, these expenditures only include packaged software, and thus form a lower bound on the total cost share for software. Because IDC began estimating software expenditure by manufacturers in 2006, we apply the compound annual growth rate of this value between 2006 and 2012 to estimate pre-2006 values. The cost share should represent an environment free of piracy. We thus divide the base rate by 0.78, which is equal to one minus the U.S. piracy rate of 22%.

Supply chain adjustment

The final step in estimating the price consequences of global piracy is to adjust for the savings that emerging market manufacturers enjoy when their suppliers engage in software piracy. Because cost of materials (i.e., the amount paid to suppliers of intermediate goods) is approximately 50% of manufacturers' overall costs (the exact estimate for the U.S. is 52.9% from the Census of Manufacturers), we increase our estimates of the direct cost advantage from software theft by this amount to account for savings in material costs due to software piracy in the supply chain. That is, we multiply by a 1.5 factor.

Our reported estimates sit in-between two extremes. If one removes the supply chain consequences entirely, the losses to U.S. manufacturers would be about two-thirds of those we report below (i.e., removing the 1.5 factor lowers the impact by 1/1.5=2/3). This extreme scenario would occur if intermediate goods were freely traded globally, such that U.S. manufacturers also saved on material costs due to unfair practices in overseas supply chains. On the other hand, our results will underestimate the harm to the extent that multiple layers of supply chains exist and provide unfair advantage (i.e., the suppliers to the suppliers of a given firm are engaging in software theft). As an upper bound, an extended supply chain could further increase our estimates by approximately one third. The mathematical maximum from an infinite series of 1.5 supply chain factors is two.

Elasticity of economic impact variables with respect to price

We model the cost savings gained by manufacturers from software piracy as a shock to the global price level for the manufactured good in question. First, we estimate the elasticity of economic variables such as employment and values of shipments and services with respect to global prices. By using the global price level, we account for the fact that unfair cost savings through piracy affect competition both in the U.S. and for global exports by U.S. firms.

Our baseline estimates of the elasticity of economic impact variables with respect to price rely on dynamic panel estimation first proposed by Arellano and Bond (1991)[1]. Arellano-Bond estimation

---

[1] Manuel Arellano and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, April 1991: pp. 277–297.

uses first differencing and previous lagged values of dependent variables as instruments to estimate the equation:

$$y_{it} = \alpha + \beta y_{i,t-1} + \delta_1 p_{it} + \delta_2 p_{i,t-1} + \rho_i + \mu_t + \varepsilon_{it}$$

We observe our panel data on U.S. manufacturing annually from 2000-2008 (year t) for each 6-digit NAICS code (industry i). In the estimating equation, ρ_i is a vector industry fixed effects that remove permanent differences in outcomes and prices across industries. Similarly, μ_t is a vector of time fixed effects that remove changes over time common for all industries (e.g., inflation). The estimation is thus focus attention on relative changes in prices and outcomes for industries.

We estimate the equation separately for three dependent variables, obtained from the NBER CES Manufacturing Industry Database[2]: log employment, log value of shipments (used to model manufacturers' revenue) and log value added (used to model manufacturers' contribution to GDP). We obtain log price p from price indices released by the Bureau of Economic Analysis.[3] As expected, we obtain positive and statistically significant elasticities of the three impact variables with respect to price:

Estimated elasticity w.r.t. price (std. errors in parentheses)

Employment	Value added	Value of shipments
1.454	1.625	1.842
(0.591)	(0.611)	(0.605)

In determining statistical significance, we cluster standard errors by industry. The interpretation of this elasticity is that a 10% increase in price for an industry is correlated with a 14.5% increase in employment.

The dynamic component of our model allows us to capture the possibility that manufacturers may take time to respond to changes in global prices. In our baseline equation, we include a single lagged value of price. We find that including a second lag increases the values of all three estimated elasticities. Similarly, the estimated magnitudes of the elasticities decrease somewhat when we estimate a static model, though they remain statistically significant and economically important. The static estimates for employment, value added and value of shipments are 0.541 (s.e. = 0.141) 1.145 (s.e. = 0.094), and 1.077 (s.e. = 0.047), respectively.

We have also found robust elasticities when looking at many other variants: For example, weighting industries based upon various measures of industry size, including detailed controls for the evolution of sub-sectors within manufacturing, examining price increases vs declines, considering price deflators, and more. It should be recognized the price can play a direct role in changing value

[2] http://www.nber.org/nberces/
[3] http://www.bea.gov/national/Index.htm

of shipments and value added since is it an underlying component, but its role for employment comes purely through quantities.

Economic impact variables

Because the quantities we estimate above are in percentage terms, we multiply by the total U.S. value for the economic impact variable in question to obtain the impact of global piracy for a given year. We obtain annual data series on employment and value added (used to model GDP) from the Bureau of Economic Analysis, and value of shipments from the U.S. Census. As our data series for value of shipments and value added are in nominal terms, our estimates reflect manufacturers' losses in terms of dollars from each year in question. If measured in current dollars, these estimated losses would increase.

Impact of global piracy on patents

By reducing the size of the market available to U.S. manufacturers, global software theft undermines the incentives of U.S. manufacturers to innovate. Researchers have established empirical relationships between market size and innovative outcomes such as patents[4], under the theory that larger markets will provide more incentives for innovators to capture greater gains.

We estimate the impact of global piracy on U.S. manufacturing patents in a manner similar to the impact on employment and revenue discussed above, except that we estimate the elasticity of patents with respect to market size rather than price. We regress log patents granted by the USPTO[5] to U.S. firms by year and 6-digit NAICS industry on log market size[6] for the industry five years prior. Our baseline estimate of the elasticity of patent grants with respect to market size of 0.382 (s.e. = 0.198) is positive and statistically significant.[7]

After estimating the elasticity, we estimate annual patents lost by U.S. manufacturers as:

$$Cost_t^{patent} = Elasticity\ w.r.t. market\ size * Piracy\ rate_t * Software\ cost_t * US\ manufacturing\ patents_t$$

The annual piracy rate and software cost as a percentage of total cost are identical to the calculations above. U.S. manufacturing patents equal total patents granted by the USPTO, times 0.94 to reflect the Patent Office's estimate that 94% of patents granted are awarded to manufacturing firms.

---

4 See, for example: Jacob Schmookler, *Invention and Economic Growth* (Cambridge, MA: Harvard University Press, 1966) and Daron Acemoglu and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry", *Quarterly Journal of Economics*, 119(3), August 2004: pp. 1049–1090.

5 http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm

6 We estimate market size as the sum of the value of global exports by NAICS 6-digit industry across countries.

7 Our baseline estimates weight industries by market size and eliminate data points that lie outside the 5%/95% range of the sample. Without weighting, we estimate an elasticity of 0.205 (s.e. = 0.068).

Tax revenue calculation

To calculate lost federal tax revenue, we first observe federal taxes (excluding payroll tax and dividend income) paid by U.S. manufacturers as a percentage of total revenue. The most recent data we observe are from 2001-2003, where taxes were on average 2.8% of value of shipments (revenue). We then multiply 2.8% by our estimate of lost manufacturers' revenue each year to obtain an annual estimate for lost federal taxes.

Opportunity for growth if global software piracy reduced 2.5% per year for years

We estimate growth in the U.S. manufacturing variables in question in a manner similar to the estimates of cost that we describe above. The piracy reduction term equals 2.5% in year 1, 5% in year 2, 7.5% in year 3 and 10% in year 4 of the simulation. We hold the economic impact variable and software cost estimates constant at their 2012 values for each of the four years in the simulation.