PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES

Whoops! Berkman Study Shows "Open Access" Reduces Broadband Consumption

George S. Ford, PhD*

November 12, 2009

Introduction

In July 2009, the Federal Communications Commission ("FCC") enlisted the Berkman Center for Internet and Society at Harvard University to "conduct an independent expert review of existing literature and studies about broadband deployment and usage throughout the world."1 FCC Chairman Julius Genachowski hoped that the review would "lay the foundation for enlightened, data-driven decisionmaking."2 Blair Levin, Coordinator of the FCC's National Broadband Plan, was looking for a "comprehensive assessment" of the literature.³ A draft of the report, Next Generation Connectivity: A Review of Broadband Internet Transitions and Policy from Around the World, was posted for comment on the FCC's website in October 2009.4

By its own account, the "most surprising and significant finding" of the Berkman Study

... is that "open access" policies – unbundling, bitstream access, collocation requirements, wholesaling, and/or functional separation – are almost universally understood as having played a core role in the first generation transition to broadband in most of the high performing countries; that they now play a core role in planning for the next generation transition; and that the positive impact of such policies is strongly supported by the evidence of the first generation broadband transition (at 11)." No doubt, this conclusion will be the target of much of the commentary on the Berkman Study, both positive and negative. In this PERSPECTIVE, I take no position on the desirability of unbundling policies in a post-POTS environment. Rather, my interest is solely in whether the economic and statistical analysis contained the Berkman Study can withstand professional scrutiny so that policymakers can rely upon its findings with confidence. Regrettably, the answer is no.

While the study's authors verbally conclude that open access policies stimulate increased consumption of broadband, the econometric model they rely upon shows the opposite – open access reduces the consumption of broadband.

As shown below, the Berkman Study first improperly estimates its econometric model and then incorrectly interprets the results from it. The error in the interpretation is significant. While the study's authors verbally conclude that open access policies stimulate increased consumption of broadband, the econometric model they rely upon shows the opposite – open access *reduces* the consumption of broadband. As shown here, the Berkman Study's authors are befuddled by their own modeling effort. Accordingly, policymakers would be remiss to accord the Berkman Study any probative weight, particularly with regard to the positive or negative effects of unbundling policies in a post-POTS world.

Reviewing the Econometric Evidence

Because the Berkman Study (at 78-80) finds the qualitative evidence for "open access" to be equivocal, it decided to supplement its qualitative discussions with a "re-analysis" of the econometric model contained in an earlier study by John de Ridder, a researcher at the OECD, entitled Catching Up in Broadband – What Will It Take? released in 2007.5 Thus, the Berkman Center relies heavily, if not exclusively, on this study for its support of open As shown in this PERSPECTIVE, this access. reliance is unfortunate, since the de Ridder study is replete with many significant (and apparent) econometric errors and its results are inconsistent with basic economic theory. Further, de Ridder's results, like Berkman's, reject the hypothesis that unbundling improves broadband consumption, though both studies conclude the opposite (since apparently neither understands the underlying economic and econometric models used).

To begin, consider the econometric model set forth in de Ridder and later adopted bythe Berkman Study. de Ridder's (2007, at 5) model is written as follows:

$$Q_{\rm D} = f(P, Y, A, E, S, W)$$
 (1)

$$Q_{\rm S} = f(\mathbf{P}, \mathbf{U}, \mathbf{G}, \mathbf{C}) \tag{2}$$

$$Q_{\rm D} = Q_{\rm S} \tag{3}$$

where Q_D is defined as the demand for broadband, Q_S is the supply of broadband, P is price and appears in both equations, Y is income, A is age, E is education, S is the addressable market, W is weather, U is population density, G is government policy, and C is competition.⁶ All of these variables are eventually assigned particular values from available data. For example, the education variable E is defined to be the percent of population age 25-64 with a tertiary education (named ETERT in the paper).

This static supply and demand model will be familiar to anyone who has taken either a basic economics or econometrics course. It is found in nearly every principles of econometrics text, including, but not limited to, Theil (1971); Maddala (1977); Pindyck and Rubinfeld (1991); Studenmund (1992); Gujarati (1995); Hill, Griffiths, and Judge (2001); and Asteriou and Hall (2007).⁷ In that regard, the model is standard fare.

Figure 1 illustrates this supply and demand framework in the familiar graphical form. There is a demand curve, labeled D, which is downward sloping. The law of demand states that quantity demanded is inversely related to price.⁸ The supply curve, labeled S, is upward sloping, since the law of supply says that quantity supplied is positively related to the price.⁹ The intersection of the two curves is the equilibrium, with equilibrium price-quantity pair (P*, Q*) in the figure. A change in a nonprice determinant of Q_D or Q_S is illustrated by a shift in the curves, whereas price changes are reflected in movement along the curves. For example, if income rises, then the demand curve shifts up and to the right (if the good is a normal good), increasing equilibrium price and Within this supply-and-demand quantity. framework, de Ridder and the Berkman Study set out to quantify the relationships between broadband connections per capita (the Q's in the model) and price (P), unbundling policies (G), and the other factors.

Why does this supply-demand model so consistently appear in basic econometrics texts? First, the supply-demand framework is understood by everyone with even a rudimentary knowledge of economics, so it is

PERSPECTIVES

familiar. Second, the supply and demand model is used to teach a very basic lesson in econometrics; specifically, Equations (1) and/or (2) should not be estimated individually by Ordinary Least Squares regression ("OLS"). These equations make up a Simultaneous Equations Model ("SEM") and must be estimated taking into account the fact that the Q and P variables are jointly determined.



As observed by Gujarati (1995), in his classic text BASIC ECONOMETRICS, 10

Now it is not too difficult to see the P and Q are jointly dependent variables.... Therefore, a regression of Q on P ... would violate an important assumption of the classical linear regression model, namely, the assumption of no correlation between the explanatory variable(s) and the disturbance term (at 637-8).

The correlation between the explanatory variable(s) and the disturbance terms, according to Hill, George and Judge (2001),¹¹

leads to failure of the least squares estimator, ... making the least squares estimator biased and inconsistent (at 306).

In other words, estimating either or both Equations (1) and (2) by OLS results in biased and inconsistent estimates of the coefficients, which means the coefficients are not good approximations, even in large samples, of their true values.

In blatant disregard for basic econometric principles, the Berkman Center estimates the equations using OLS, thereby committing the grossest of econometric errors.

In blatant disregard for basic econometric principles, de Ridder and the Berkman Study estimate the equations using OLS, thereby committing the grossest of econometric errors.¹² The use of OLS in this context is explicitly rejected every basic and advanced in econometrics textbook.¹³ Proper technique requires Equations (1) and (2) to be estimated using a procedure that accounts for the joint estimation of P and Q, such as Two Stage Least Squares ("2SLS"), which is a commonly used technique for SEMs.14

To illustrate the defects with the econometric model in the Berkman Study, I first discuss and replicate the estimates from the de Ridder and Berkman studies. In doing show, I reveal why the Berkman Study is precisely backwards in its conclusion on the effects of unbundling. According to the estimates in the Berkman Study, unbundling *reduces*—not increases—the equilibrium number of broadband connections per capita. The false conclusion arises from the fact that the Berkman Center does not understand its own model. Notably, my analysis is based on their own estimates, and does not reflect any modifications to their work.

Then, I estimate the models using 2SLS to illustrate the impact of poor econometric technique. This re-estimation effort merely exposes one of the many severe defects in the Berkman Study's analysis of unbundling. Proper estimation is not, however, an elixir for all of the ills that plague the Berkman Study. 2SLS is shown to have a significant effect on the outcomes of interest, though, in the end, the conclusions (when properly drawn from the evidence) are not reversed—unbundling policies either reduce or have no effect on broadband consumption.

The Effect of Unbundling Using the Berkman Model

Final specifications of the SEM proposed by de Ridder and replicated by Berkman include the Demand Curve:

$$QTOT = \alpha_0 + \alpha_1 ln(PDSL) + \alpha_2 YINDEX \qquad (4) + \alpha_3 AGE + \alpha_4 ETERT + \alpha_5 SIP + \alpha_6 SUN + \varepsilon$$

where QTOT is broadband connections per ln(PDSL) log capita, is the natural transformation of the price of a DSL connection, YINDEX is a measure of per-capita GDP, AGE is the share of total population aged 35-44, ETERT is the percent of the population aged 25-64 with a tertiary education, SIP is total Internet subscriptions as a percent of population, SUN is the daily average number of sunlight hours, and ε is the econometric disturbance term; and the Supply Curve:

$$QTOT = \beta_0 + \beta_1 ln(PDSL) + \beta_2 UURB$$
(5)
+ \beta_3 CFAC + \beta_4 GUYRS + \vee v

where QTOT is broadband connections per capita, UURB is the percent of urbanized population, CFAC is the share of non-DSL lines in QTOT, GUYRS is the number of years a local loop unbundling policy has been in place, and v is the econometric disturbance term for this equation. (I will address the legitimacy of the SIP, GUYRS, and CFAC variables later, but I suspend criticism for the moment.) Given the law of demand, we expect α_1 to be negative ($\alpha_1 < 0$), and from the law of supply we expect β_1 to be positive ($\beta_1 > 0$), rendering the standard relationships illustrated in Figure 1.

The OLS estimates using the 2005 data from de Ridder and the Berkman Study (30 observations) can be summarized as:

Demand Curve:

$$QTOT = -1.865 \cdot \ln(PDSL) + AX + e$$
 (6)

Supply Curve:

$$QTOT = -2.79 \cdot \ln(PDSL) + 0.57 \cdot GUYRS$$
 (7)
 $+ BZ + v$

where AX and BZ represent the other covariates in the model and their coefficients.¹⁵ The positive coefficient on GUYRS ($\beta_4 = 0.57$) leads de Ridder to conclude that unbundling positively effects broadband consumption, and the same conclusion is made in the Berkman Study.

[Under] the Berkman Study's analysis] the supply curve is downward sloping! This result implies that as broadband prices rise, network operators supply less broadband. Intuitively, this result makes little sense, violates the law supply, and muddles of interpretation.

Problems with the de Ridder and Berkman analysis are as obvious as numerous. First, the two equations are estimated by OLS in both studies, which is inappropriate under their own model specification. Second, the coefficient on PDSL is the supply equation is negative. *The supply curve is downward sloping!* This result implies that as broadband prices rise, network operators supply less broadband. Intuitively, this result makes little sense, violates the law of supply, and muddles interpretation.¹⁶ (If a skilled economist or econometrician observed these results, he or she would conclude that the supply curve was not identified by the model

PERSPECTIVES

and conduct further analysis. But, de Ridder and the Berkman Study proceed as if all is well.)

The negative sign on price in the supply curve is not a trivial matter, nor just an econometric problem, as the negative sign leads both studies to bungle the interpretation of the econometric results. Indeed, rather than indicating the "open access" policies are good for broadband adoption, both the de Ridder and Berkman studies' econometric models indicate that unbundling reduces broadband consumption. Let me explain.

Should policymakers believe that unbundling reduces broadband consumption? Certainly not based on the de Ridder and the Berkman studies. The economic and econometric analysis in both are clumsily conducted and incorrectly interpreted, rendering them unsuitable for use in formulating public policy.

Figure 2 illustrates the implication of the econometric estimates of the de Ridder and the Berkman studies. In the absence of unbundling, the demand curve is labeled D_0 , and the supply curve is labeled S_0 . As shown, both the supply and demand curves are downward sloping as indicated by the econometric estimates from these studies (see Eqs. 6 and 7). Also, the econometric estimates imply the supply curve is flatter (more elastic) than the demand curve, as is illustrated in the figure (also see Eqs. 6 and 7).¹⁷ The "no unbundling" equilibrium is defined by the intersection of the supply and demand curves at values (P_0 , Q_0) as illustrated.¹⁸



In both de Ridder and the Berkman Study, the argument is made that unbundling shifts the supply curve, since the coefficient on the GUYRS variable is positive and statistically significant in some models. This is correct. A positive coefficient on GUYRS in the supply equation means that quantity supplied of broadband connections is larger at all prices under an unbundling regime. As such, the effect of unbundling is illustrated in Figure 2 by a shift in supply from S_0 to S_1 . After the shift in other things constant, supply, the new equilibrium is defined to be (P_1, Q_1) . As shown in the figure, according to the econometric results in both de Ridder and the Berkman Study, unbundling increases supply, and in doing so, increases equilibrium price and reduces the equilibrium quantity of broadband connections per capita. This perverse effect, illustrated in Figure 2, arises because the supply has a negative slope and a larger price elasticity than does the demand curve.

The coefficients in Equation (6) and (7) are directly from de Ridder (and confirmed by the Berkman Study) and based on 2005 data for the 30 OECD members. Berkman focuses more on the panel data (adding year 2002 to the 2005 data) used by de Ridder. The estimates of interest using the panel data (54 observations) are: Demand Curve:

 $QTOT = -1.5 \cdot \ln(PDSL) + AX + e$ (8)

Supply Curve:

 $QTOT = -1.9 \cdot \ln(PDSL) + 0.77 \cdot GUYRS + (9)$ BZ + v

As shown in Equations (8) and (9), there is no change in the interpretation of the outcomes shown in Figure 2; the supply curve remains downward sloping and more elastic than the demand curve, and the coefficient on GUYRS is positive. Therefore, the conclusion is retained: unbundling raises equilibrium prices and reduces the consumption of broadband under the de Ridder and Berkman Study model. The larger is the coefficient on GUYRS, the greater is the price increase and the greater is the quantity reduction.

It is possible to compute the equilibrium effect of unbundling policies shown in Figure 2 given the estimated coefficients of the models. For the 2005 data (n = 30), the equilibrium quantity of broadband connections per capita falls by about 8% for each additional year of unbundling.¹⁹ For the panel data (n = 54), Q* falls by about 16% for each additional year of unbundling.²⁰ (Keep in mind these reductions are merely artifacts of a poor econometric model.)

Should policymakers believe that unbundling reduces broadband consumption? Certainly not based on the de Ridder and the Berkman studies. The economic and econometric models in both are improperly conducted and incorrectly interpreted, rendering them unsuitable for use in formulating public policy.

Appropriate Estimation Methods

As noted above, the model proposed by de Ridder and adopted by the Berkman Study is a Simultaneous Equations Model. Such a model should not be estimated by OLS; special estimation techniques are required. In this section, I re-estimate the models using an accepted technique—Two Stage Least Squares ("2SLS"). The focus of the Berkman Study is on Equation (5) above, so I will limit my attention to that equation in this section. Again, this analysis does not suggest that proper estimation techniques make the model valid. They do not. The modeling specification and the data used remain inescapably unsuitable to quantify the effects of unbundling in either the POTS or post-POTS era.

First, I replicate the OLS analysis using the 2005 data employed in both studies.²¹ The estimated coefficients and their t-statistics (both standard and robust) are provided in Table 1, and are consistent with those of reported by the de Ridder and Berkman studies. Since the equation is estimated with only 30 observations, I also provide bootstrapped (robust) t-statistics in brackets to account for the small sample. The bootstrapped t-statistics are not much different from their robust values.

Table 1. Replication (OLS)				
Coef.	Estimate	t-stat (robust) [bootstrapped]		
Constant: β ₀	-5.36	-0.77 (-0.75) [-0.75]		
$ln(PDSL): \beta_1$	-2.79	-3.76 (-4.74) [-3.55]		
UURB: β_2	0.32	4.10 (4.61) [4.41]		
CFAC: β_3	0.06	1.11 (1.15) [1.15]		
GUYRS: β_4	0.57	1.65 (1.23) [1.23]		

As shown in Table 1, the coefficient on price is negative and statistically significant at standard levels (meaning the supply curve slopes downward), the coefficient on UURB is positive and statistically significant, and the coefficient on CFAC is positive but not statistically different from zero. The coefficient on GUYRS is positive and achieves a minimal level of statistical significance (just short of 10%), but only if nonrobust t-statistics are used. For the robust and bootstrapped t-statistics, **GUYRS** is not statistically different from zero at anything near standard levels. (Berkman selectively uses robust t-statistics (at 139), but not in this case.) Generally speaking, good practice calls for the use of robust standard errors.²² Note that the

coefficient estimates in Table 1 are invalid since they are estimated by OLS.

Table 2 provides the 2SLS estimates of the equation, which is a proper estimation approach for the specified model and provides consistent estimates of the parameters.²³ The estimated coefficients change dramatically in some instances. The coefficient on the price variable (which is now treated as endogenous) has become larger (in absolute value), changing from -2.79 to -4.89. This change is expected, since simultaneity bias in OLS tends to drive the coefficients toward zero. The coefficient on GUYRS is now very small and nowhere near statistically significant (with a t-statistic of about 0.02). The other coefficients and t-statistics are not much changed. Overall, the differences methodologies across strongly indicate simultaneity bias. Moreover, estimates from proper methodology do not support the conclusion that unbundling matters for broadband consumption. The supply curve remains downward sloping and more elastic than the demand curve.²⁴ This suggests that significant specification errors still remain in the Authors' attempts to portray international broadband markets.

Table 2. Results				
(2SLS; 2005 data)				
Coef.	Estimate	t-stat (robust)		
		[bootstrapped]		
Constant: β_0	5.86	0.57 (0.59) [0.51]		
ln(PDSL): β_1	-4.99	-3.10 (-2.83) [-2.50]		
UURB: β ₂	0.29	3.47 (4.24) [3.83]		
CFAC: β_3	0.07	1.23 (1.28) [1.05]		

0.01

An interesting and important question is whether from a statistical standpoint we can treat ln(PDSL) as exogenous and produce somewhat reliable results using OLS?²⁵ For that answer, I turn to the standard approach for testing exogeneity and perform a Hausman test.²⁶ For the 2005 data, the null of that test exogeneity of ln(PDSL)—is rejected at the 10% level. So, not only is the model is explicitly

0.03 (0.02) [0.02]

specified as a SEM, it should be estimated as a SEM.

While the Berkman Study mentions the potential for simultaneity at one point, it never considers the simultaneity with Q and P, despite the explicit specification of a SEM in de Ridder. While the Berkman Study does contemplate the endogeneity of PDSL and GUYRS (at 116), the analysis is inept. First, the proper consideration is the endogeneity of PDSL and/or GUYRS with QTOT, not with each other. Saying that two regressors may be endogenous with each other is senseless; endogeneity involves relationships dependent (or among the endogenous) variables. Second, the Berkman Study claims to evaluate simultaneity by regressing one variable on the other (PDSL and GUYRS), and then concluding because the R² is not equal to 1.0 the two variables are only "partially endogenous (at 116)." The analysis is asinine. Correlation is not a test of exogeneity (or endogeneity), and determined simultaneity cannot be bv evaluating relationships between two righthand side variables. (The Berkman Study has confused multicollinearity with endogeneity.) Third, their proposed solution to simultaneity is bizarre, choosing simply to run the regression "without the partially endogenous factor of DSL price (at 116)." While there are legitimate means of dealing with an endogenous variable (e.g., 2SLS), tossing the variable out and estimating the regression is not one of them.²⁷ Moreover, the Berkman Study proposes to estimate a supply curve without a price variable. The Berkman Study's entire analysis on endogeneity is, put bluntly, inept.

Like de Ridder, the Berkman Center also adds 2002 data to the sample. The panel data increases the statistical significance of the GUYRS variable in their improperly estimated models.²⁸ This result is unsurprising given that the specification of the GUYRS variable acts much like a dummy variable capturing, in part, the trend in the data. Berkman makes no effort to remedy this obvious consequence of their

GUYRS: β_4

specification. As shown above, given the entire model, the larger coefficient and higher statistical significance of GUYRS only provides greater certainty that unbundling policies reduce broadband consumption.

Using the panel data, the OLS replication and 2SLS results are presented in Table 3. The variable D2005 is a dummy for the 2005 period. The GUYRS variable is defined to be nonnegative. I provide both White's robust and the cluster-robust standard t-statistics (in parenthesis) for consistency with Boyle et al. (2008).²⁹ (Notably, the proper estimation of the standard errors has no impact on the conclusion unbundling reduces broadband that consumption in these models.)

Table 3. Results (OLS and 2SLS; 2002 and 2005 data)				
		(Cluster Robust)		
OLS				
Constant: β_0	-11.30	-2.19 (-1.83)		
$ln(PDSL): \beta_1$	-1.88	-3.52 (-3.12)		
UURB: β_2	0.26	5.21 (4.27)		
CFAC: β_3	0.04	1.68 (1.40)		
GUYRS: β ₄	0.76	2.32 (1.96)		
D2005: β ₅	7.15	5.20 (5.64)		
2SLS				
Constant: β_0	-1.60	-0.15 (-0.12)		
$\ln(PDSL): \beta_1$	-3.63	-2.41 (-1.92)		
UURB: β ₂	0.24	3.97 (3.18)		
CFAC: β_3	0.05	1.82 (1.57)		
GUYRS: β ₄	0.33	0.66 (0.55)		
D2005: β ₅	6.11	3.47 (3.35)		

As with the 2005 data, the influence of simultaneity bias is seen clearly in the estimates: the coefficient on price is larger (in absolute value). The supply curve remains downward sloping. All other coefficients are similar with the exception of GUYRS, with both the coefficient and t-statistics shriveling under 2SLS estimation. The GUYRS variable is not anywhere near statistically significant when the equation is properly estimated as a SEM, meaning unbundling has no statistically significant effect in the supply equation, though the coefficient is positive. The supply curve

remains downward sloping and more elastic than the demand curve ($\alpha_1 = -0.95$; $\beta = -3.63$). (That is, the model is still junk science even if estimated using correct techniques.)

My conclusions here are not much affected if I apply the mixed-effects regression proposed in the Berkman Study. The supply curve remains downward sloping and of larger elasticity than the demand curve. Unbundling, then, by the model proposed in the Berkman Study, reduces broadband consumption.

Other Model Mis-Specifications

There are two other significant defects in de Ridder's model left unresolved by the Berkman Study. The first is detailed in PHOENIX CENTER POLICY PERSPECTIVE NO. 09-02, released in May 200930 (but ignored in the Berkman Study, revealing a incomplete review of the literature). In this PERSPECTIVE, I provided a detailed analysis demonstrating that de Ridder's model is mis-specified due to the treatment of the SIP variable; an error repeated in the Berkman Study. The SIP variable is defined as broadband plus dialup connections per capita. In essence, de Ridder regressed broadband adoption on broadband adoption, avoiding collinearity by adding to it dialup connections and using lags.³¹ The same error appears in a study by the Information Technology and Information Foundation, so it is a common mistake.³² As shown in PERSPECTIVE 09-02, the misspecification takes the form of an invalid coefficient constraint (assuming identical effects on broadband by broadband and dialup connections), with predictable consequences (biased coefficients). I direct the reader to that PERSPECTIVE for more details. For an alternative maturity variable, see also PHOENIX CENTER PERSPECTIVES NOS. 08-03 and 09-01.33

Second, the CFAC variable is problematic. The dependent variable of the de Ridder model is broadband connections over population. CFAC is defined as the cable connections over

broadband connections. Thus, the denominator of CFAC is the numerator of the dependent variable.³⁴ It is possible that this linkage makes CFAC endogenous to the model. The Berkman Study makes no effort to address this problem.

Improper Modifications of the Data

As noted above, the GUYRS variable used in the de Ridder and Berkman studies measures the number of years a local loop unbundling regime has been in place. For example, in de Ridder, the variable has a value of 10 for the United States, reflecting the difference between yearend 2005 and the passage of the Telecommunications Act of 1996 in February of that year. Note that the use of a "years since unbundling" implies that the effects of an unbundling policy occur with a lag; this is an important aspect of the model specification ignored by the Berkman Study.

The Berkman Center modifies the GUYRS variable as constructed by de Ridder for a number of countries, increasing and decreasing its value based on what it claims are relevant particularities of individual countries. For Australia, the GUYRS value of 6 is reduced to 1 to reflect "the high prices until the competition notice issued in March 2004 ... which reportedly led to lower prices (at 149)." Canada, similarly, has its GUYRS value reduced from 9 to 5, to reflect "initial sunset and high LLU rates (at 149)." Hungary, alternately, remains unadjusted due to "Insufficient data to form an opinion (at 150)." The U.S. value of 10 is reduced to 0 to reflect "2001-02 FCC decisions to shift to intermodal competition (at 150)."

The problems with such modifications to the data are astoundingly obvious. For example, many of the adjustments are based on the prices for local loops or other particulars of a given country. Accounting for price differential is important because it demonstrates that an unbundling regime is not a binary condition (it is nuanced, not simply on-or-off).³⁵ The

presence of an unbundling regime is irrelevant if the loop prices are too high, as Berkman observes. In the United States, for example, the number of unbundled loops purchased varied significantly across the states largely due to price differences, yet all were under the same federal unbundling regime.³⁶ Yet, neither de Ridder nor the Berkman Study include a loop (or unbundled element) price variable in their regressions, which is a gross mis-specification of the model. Trying to incorporate price information by manipulating GUYRS, as the Berkman Study does, is improper.

... throughout the Berkman Study, the authors are separating the sick rats from the well ones and then assigning the treatment ex post. This scheme is taboo among research scientists ...

The manipulation of the GUYRS variable is but one example of the Berkman Study's outcomesmotivated, and thus improper, analysis. Sound empirical research of treatments and outcomes requires the researcher to ignore the observed outcomes in formulating the hypothesis tests and choosing the empirical methodologies. Yet, the Berkman Study peeks at the outcome and then tries to formulate some procedure to attribute observed differences to one factor or In other words, throughout the another. Berkman Study, the authors are separating the sick rats from the well ones and then assigning the treatment ex post. This scheme is taboo among research scientists, since such outcomesdriven analyses are likely to render biased results, both in a statistical sense and by the introduction of researcher bias.

Second, a particularly egregious modification of GUYRS variable relates to the treatment of the United States. As noted above, Berkman sets the

PERSPECTIVES

modified value of the GUYRS variable to zero for the United States, claiming that unbundling ended in the United States in 2001-02. There are (at least) two problems with this change to the data. First, it is entirely inconsistent with the use the GUYRS variable. The specification of GUYRS—years since unbundling regime was in place—implies a lagged effect of unbundling on outcomes. Even if unbundling had stopped in 2002—it did not—setting the U.S. value to zero is improper, since it explicitly rejects a lagged effect of unbundling. The modification to GUYRS for the United States lacks internal consistency.



Second, the modification conflicts plainly with the facts. In Figure 3, both OECD and FCC data on unbundled loops in the U.S. are illustrated. In contrast to the claim in the Berkman Study, the number of unbundled loops in the U.S. peaked in 2004 and 2005, increasing by two-fold from 2001. Also, according to OECD data, of the countries reporting data in 2005 the U.S. had more unbundled loops as a percentage of access paths than any other OECD member country except Iceland.³⁷ The facts (and theory) plainly reject the Berkman Study's modification of the GUYRS variable.

Conclusion

The intended purpose of the Berkman Study is to aid the FCC in establishing a National Broadband Plan by reviewing the broadband agendas of various countries and evaluating the existing research on broadband efforts. Ideally, some useful guidance, or a useful review of the existing evidence, would be provided. At these tasks the study fails miserably.

The Berkman Study provides a woefully inadequate review of the literature and offers no new data, no new methods, and no innovative policy ideas. The statistical analysis is mostly unskilled and unenlightening. Indeed, the economic and econometric analysis used in the Berkman Study to support its "most significant finding" that unbundling improves broadband consumption is embarrassingly bad. The analysis is so convoluted that even the Study's authors cannot understand the results. The Berkman Study claims that "open access" stimulates broadband consumption, but the correct interpretation of its own evidence is that unbundling reduces broadband consumption.

Chairman Genachowski's desire for "enlightened, data-driven decisionmaking" is not served well by the Berkman Study. Accordingly, given the multitude of technical flaws outlined herein, how much credibility the FCC accords the Berkman Study in the end will provide a clear bellwether of the Commission's commitment to the intellectual rigor it purports to want.

P E R S P E C T I V E S

NOTES:

* Dr. George Ford is Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies. The views expressed in this PERSPECTIVE do not represent the views of the Phoenix Center, its Adjunct Follows, or any if its individual Editorial Advisory Board Members.

¹ <u>http://www.fcc.gov/Daily_Releases/Daily_Business/2009/db0714/DOC-291986A1.pdf</u>.

² Id.

³ Id.

⁴ <u>http://www.fcc.gov/stage/pdf/Berkman_Center_Broadband_Study_13Oct09.pdf.</u>

⁵ J. de Ridder, *Catching Up in Broadband – What Will It Take*, Directorate For Science, Technology and Industry, Committee for Information, Computer and Communications Policy, Organization for Economic Cooperation and Development (2007) (available at: <u>http://ideas.repec.org/p/oec/stiaab/133-en.html</u>).

⁶ The model is formally defined at de Ritter (2007), *id.* at 5.

⁷ H. Theil, PRINCIPLES OF ECONOMETRICS (1971) at Ch. 9; G. S. Maddala, ECONOMETRICS (1977) at Ch. 11; R. Pindyck and D. Rubinfeld. ECONOMETRIC MODELS & ECONOMIC FORECASTS (1991) at Ch. 11; A. H. Studenmund, USING ECONOMETRICS (1992) at Ch. 14; and D. Gujarati, BASIC ECONOMETRICS (1995) at Ch. 18; C. Hill, W. Griffiths, and G. Judge, UNDERGRADUATE ECONOMETRICS (2001) at Ch. 14; D. Asteriou and S. Hall, APPLIED ECONOMETRICS (2007) at Ch. 11.

⁸ See, e.g., R. Ekelund Jr. and R. Tollison, ECONOMICS (1994) at 86.

⁹ Id. at 93.

¹⁰ Supra n. 7.

¹¹ *Supra* n. 7.

¹² The Berkman Study's use of mixed effects regression does not resolve the simultaneity problem.

¹³ Estimation by OLS is a common criticism of single equation models in which there is a plausible case (and in some case an implausible case) for simultaneity of one or more of the right-hand side variables. One may argue, therefore, that my pointing this out is simply a quibble over specification. Under some circumstances, I would agree; but not here. de Ridder, and subsequently the Berkman Study, explicitly specify a simultaneous system but then ignore this simultaneity in estimation. As such, my comments are not directed at a disagreement over specification, but rather simply pointing out that given their own specification, they have failed to estimate the model appropriately.

¹⁴ See, e.g., Gujarati (1995), supra n. 7, at Ch. 18.

¹⁵ de Ridder (2007), *supra* n. 5, at Table 4. An alternative price variable is LNPTOT, which de Ridder uses in some specifications of the demand curve. I use PDSL since that variable can be used in both equations as specified by de Ridder, and there is panel data available on that price. I will not address the issue of specifying a SEM and using different prices (which is senseless).

¹⁶ A downward-sloping at supply curve is not necessarily prohibited by theory. However, equilibrium may not be stable with downward sloping supply curves, particularly in the case considered here when the supply curve cuts the demand curve from below.

¹⁷ The coefficient on price measures $\Delta Q/\Delta P$, so that larger values of the coefficient imply a larger elasticity (the elasticity of demand is computed using [$(\Delta Q/\Delta P)(P/Q)$].

¹⁸ I reiterate that the equilibrium is unstable since the supply curve cuts the demand curve from below.

¹⁹ Roughly, the demand curve is $Q_D = 20 - 1.9P$, where the first term is AX computed at the means of X, and the supply curve is $Q_S = 20 - 2.8P + 0.57G$, where the first term is BX computed at the sample means and excludes G. The equilibrium price is: $P^* = 0.6G$. Inserting P* into either the demand or supply curve renders $dQ^*/dG = -1.1$ on a mean of QTOT of 14.14.

NOTES CONTINUED:

²⁰ *Id.* $(dQ^*/dG = -1.65 \text{ on a mean of QTOT of 9.8}).$

²¹ I am grateful to the authors of Boyle *et al.* (2008) and Y. Benkler at The Berkman Center for sharing their data. G. Boyle, B. Howell, and W. Zhang, *Catching Up in Broadband Regressions: Does Local Loop Unbundling Really Lead to Material Increases in OECD Broadband Uptake?* (July 28, 2008) (available at: <u>http://ssrn.com/abstract=1184339</u>). Following the Berkman Study, the GUYRS variable is modified to have only non-negative values.

²² T. Wallace and J. Silver, ECONOMETRICS: AN INTRODUCTION (1988) ("Generally speaking, it is probably a good idea to use the WHITE option ... routinely (at 265)").

²³ 2SLS estimates are biased but consistent. That is, they converge to their true values as the sample size gets larger. The instruments are those found in de Ridder and the Berkman Study: YINDEX, AGE, ETERT, SIP, SUN, UURB, CFAC, and GUYRS.

²⁴ The 2SLS estimate of the coefficient on price for the demand curve is -0.83.

²⁵ In a richer model that is not formulated in the supply-demand framework, price may be legitimately treated as exogenous, though statistical evaluation of exogeneity is good practice. *See, e.g.,* G. S. Ford, T. M. Koutsky and L. J. Spiwak, *The Broadband Performance Index: A Policy-Relevant Method of Comparing Broadband Adoption Among Countries,* PHOENIX CENTER POLICY PAPER NO. 29 (July 2007) (available at: <u>http://www.phoenix-center.org/pcpp/PCPP29Final.pdf</u>) (addressing endogeneity of price at 23).

²⁶ Gujarati (1995), *supra* n. 7, at 670-2.

²⁷ Standard procedures, such as 2SLS or reduced form regression, address the problem by replacing the endogenous variable with another variable or set of variables.

²⁸ Boyle *et al.* (2008), *supra* n. 21, re-estimate de Ridder's panel data model and account for the clustering as discussed in White (1984), Moulton (1986), Arellano (1987), and Wooldridge (2003). H. White, ASYMPTOTIC THEORY FOR ECONOMETRICIANS (1984); Moulton, B.R. (1986), *Random Group Effects and the Precision of Regression Estimates*, 32 JOURNAL OF ECONOMETRICS 385-397 (1986); M. Arellano, *Computing Robust Standard Errors for Within-Group Estimators*, 49 OXFORD BULLETIN OF ECONOMICS AND STATISTICS 431-434 (1987); M. Arellano, *Computing Robust Standard Errors for Within-Group Estimators*, 49 OXFORD BULLETIN OF ECONOMICS AND STATISTICS 431-434 (1987); J. Wooldridge, *Cluster-Sample Methods in Applied Econometrics*, 93 AMERICAN ECONOMIC REVIEW 133-138 (2003). In response to Boyle *et al.* (2008), Berkman argues that the cluster technique is non-standard and relatively new (at 139). This response is ridiculous. The cluster adjustment is now standard, frequently used and available in most econometric packages. A number of refinements to the cluster techniques have been proposed over the years. *See e.g.*, G. Kauermann and R. Carroll, *A Note on the Efficiency of Sandwich Covariance Matrix Estimation*, 96 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 1387-1396 (2001); L. Mancl and T. DeRouen, *A Covariance Estimator for GEE with Improved Finite- Sample Properties*, 57 BIOMETRICS 126-134 (2001). Furthermore, the approach is detailed in many textbooks. *See, e.g.*, M. Verbeek, A GUIDE TO MODERN ECONOMETRICS (2008); C. Baum, AN INTRODUCTION TO MODERN ECONOMETRICS USING STATA (2006).

²⁹ Boyle *et al., supra* n. 21.

³⁰ G. S. Ford, *Econometric Analysis of Broadband Subscriptions: A Note on Specification*, PHOENIX CENTER PERSPECTIVES NO. 09-02 (May 12, 2009)(available at: <u>http://www.phoenix-center.org/perspectives/Perspective09-02Final.pdf</u>).

³¹ *See id.* for a discussion of using lagged values.

³² R. D. Atkinson, D. K. Correa, J. A. Hedlund, *Explaining International Broadband Leadership*, Information Technology and Innovation Foundation (May 2008) (available at: <u>http://www.itif.org/files/ExplainingBBLeadership.pdf</u>).

³³ G. S. Ford, *Normalizing Broadband Connections*, PHOENIX CENTER PERSPECTIVES NO. 09-01 (May 12, 2009)(available at: http://www.phoenix-center.org/perspectives/Perspective09-01Final.pdf); G. S. Ford, *Broadband Expectations and the Convergence of Ranks*, PHOENIX CENTER PERSPECTIVES NO. 08-03 (Second Edition) (October 1, 2008) (available at: http://www.phoenix-center.org/perspectives/Perspective08-03Final.pdf).

³⁴ That is, y/x is regressed on z/y.

NOTES CONTINUED:

35 The idiosyncrasies of unbundling regimes suggest (to me at least) that there is no proper econometric specification of an empirical model aimed to test the broad effects of unbundling in an international setting. The set of variables required to accurately capture the nature of an unbundling regime are too numerous to specify in an econometric model of international outcomes, and some relevant factors are likely to be collinear with a country's fixed effect. Put simply, it is impossible to accurately measure the broad impacts of unbundling regimes in an international context using regression analysis. In POLICY PAPERS NO. 29 and 33, we relegate unbundling regimes to the disturbance term due to their idiosyncratic nature. The importance of nuances can be significant. My own research demonstrated that the four-line rule for unbundled switching significantly reduced the number of access lines and switches operated or owned by CLECs. See G. Ford, An Empirical Examination of the Unbundled Local Switching Restriction, Z-TEL POLICY PAPER NO. 3 (2002) (available at: http://www.aestudies.com/library/zpp3.pdf); G. Ford, Does Unbundling Really Discourage Facilities-Based Entry? An Econometric Examination of the Unbundled Local Switching Restriction, Z-TEL POLICY PAPER NO. 4 (2002)(available at: http://www.aestudies.com/library/zpp4.pdf). Wallsten and Hausladen attempt to determine the effects of unbundling on next-generation networks in their paper S. Wallsten and S. Hausladen, Net Neutrality, Unbundling, and their Effects on International Investment in Next-Generation Networks, 8 REVIEW OF NETWORK ECONOMICS 90-112 (2009). Their model has network deployment across countries being determined only by GDP per capita, unbundled lines per capita, and fixed effects, which is too parsimonious (in my opinion) for strong conclusions. Notably, strong conclusions are avoided, confirming only the presence of a correlation rather than the stronger claim of causation ("We find a significant negative correlation between the number of unbundled DSL connections per capita and the number of fiber connections (at 90)").

³⁶ See, e.g., T. R. Beard and G. S. Ford, Make or Buy? Unbundled Elements as Substitutes for Competitive Facilities in the Local Exchange Network, PHOENIX CENTER POLICY PAPER NO. 14 (September 2002); T. R. Beard and G. Ford, Are Unbundled and Selfsupplied Telecommunications Switching Substitutes? An Empirical Study, 12 INTERNATIONAL JOURNAL OF THE ECONOMICS OF BUSINESS 163-181 (2005); R. B. Ekelund Jr., and G. S. Ford, Preliminary Evidence on the Demand for Unbundled Elements in Telephony," 30 ATLANTIC ECONOMIC JOURNAL 443-444 (2002); A. Kline, The Demand for Unbundled Elements in Telephony Revisited, 31 ATLANTIC ECONOMIC JOURNAL 119-120 (2003); G. S. Ford and L. J. Spiwak, The Positive Effects of Unbundling on Broadband Deployment, PHOENIX CENTER POLICY PAPER NO. 19 (September 2004). See also the Local Telephone Competition reports published by the FCC (various years)(available at: http://www.fcc.gov/wcb/iatd/comp.html).

³⁷ OECD COMMUNICATIONS OUTLOOK 2009 at Tbls. 2.9 and 4.2. Data is not available for all countries.



PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES 5335 Wisconsin Avenue, NW, Suite 440 Washington, D.C. 20015 Tel: (+1) (202) 274-0235 • Fax: (+1) (202) 318-4909 www.phoenix-center.org