APPENDIX C: TECHNICAL ANALYSIS OF FIRM PRODUCTIVITY ANALYSIS TO

THE RESIDENTIAL AND COMMERCIAL BENEFITS OF RURAL BROADBAND: EVIDENCE FROM CENTRAL APPALACHIA

FINAL REPORT

JULY 2005

MARK L. BURTON

MICHAEL J. HICKS

Introduction

The final analyses provided in this study were estimates of firm productivity associated with broadband. The first of these studies was a firm level analysis, followed by regional industry level analyses and finally a firm location model (which was descriptive, not econometric). We review the first two of these models in this appendix.

Benefits to Commerical Users

To evaluate the relationship between firm productivity and broadband we estimated the following general function:

$$\frac{Y}{N} = f(\Gamma, Broadband, \Phi, \Omega)$$

Where the revenue per worker in each firm is a function of access to highway infrastructure Γ, broadband telecommunications access, and the control variables of regional characteristics Φ, and firm specific characteristics, Ω. We also permitted higher order values of some variables (to estimate non-linear relationships) and interactions between selected variables. The basic construct of the model is derived from Hicks [2001] but with addition of the broadband variable and interaction term. For readers interested in a more detailed treatment of production technologies, we recommend the earlier research.

The estimation of this model required some specific statistical assumptions which we feel are useful to highlight. First, the model was estimated using both a limited dependent variable model (LDV) and ordinary least squares (OLS) techniques. The limited dependent variable model was estimated using a statistical assumption which placed less emphasis on extreme productivity values.[1] This is important since there are potential data concerns apparent in some higher ranges of productivity estimates. Also, we excluded firms that did not have employees (e.g. sole proprietorships) and those with implausibly low levels of productivity per worker (less than $10,000 per year). This could be to data errors, or the misreporting of temporary employees as full time equivalents. These data cleaning efforts are designed to eliminate from consideration known data errors and situations where we may capture data on a very non-traditional firms in which a sole proprietor may operate the business formally from home, without employees. This type of operation may, or may not, be sensitive to broadband, but warrants separate consideration. Examples of the industries we eliminated in this process were consultancies, legal services and real estate brokers. In no instance did we eliminate all the firms in an industry, and our final sample after these selection processes was in excess of 2,700 firms. Estimation results are provided in Table 3 (page 18) of the text. Selected diagnostics accompany these results. This model was subjected to a

[1] Formally, this is the extreme value distributional assumption where we assume $f(x) = \exp(x-e^x)$.

Hausman endogeneity test, demonstrating no levels of endogeneity at any level of statistical significance.

The Broadband and Worker Wage Analyses

In estimating the impact on worker wages of broadband, we are implicitly considering the traditional models of worker productivity and workers wages in a competitive setting.[2] The empirical model construct takes the form:

$$Y_{i,t} = \alpha + \gamma_1 \Omega_{i,t} + \gamma_2 \Omega_{i,t-1} + \rho_1 W_i Y_{i,t} + \rho_2 W_i Y_{i,t-1} + \beta \, \mathrm{X}_i + \lambda Y_{i,t-1} + u_i + e_t$$

where

$$W_i = \begin{pmatrix} w_{11} & \mathrm{K} & w_{1n} \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ w_{m1} & \mathrm{L} & w_{mn} \end{pmatrix}$$

where *Y* is real incomes in county *i*, in time *t.* Following the equality operator, α denotes the common intercept, and γ the estimated coefficients for Ω the current and lagged proportion of broadband access in county *i*. The following terms ρWY denote the spatial autocorrelation function and its lagged value. The spatial autocorrelation term is the value of the dependent variable for the surrounding regions weighted by the distance to county *i*. In this example we employ W, as a first order contiguity matrix for Y. The remaining coefficients include a recession binary variable X, a 2001 dummy and the autocorrelation function for Y. The fixed effects dummy u, and the disturbance term e complete the basic form of our model. This model was estimated for selected two-digit SIC sectors individually for the states of West Virginia, Kentucky, Ohio, and Pennsylvania and collectively within a multi-state specification.[3]

[2] Virtually all macroeconomic models, and regional wage models treat the marginal productivity of labor as the labor demand curve. We do as well, believing that labor markets, especially during this time period were quite competitive at the county level.

[3] As noted in the text, in each specification, we were concerned with the possible influence of endogeneity. Accordingly we adopted the preferred correction for this problem - a two-stage least squares estimator Selection of instruments is always challenging, and in this case we employed population, real per capita

In each case we were motivated by concerns for the question of endogeneity (or direction of causation) to employ an instrumental variable technique. Our two stage least squares estimates provide a method for a consistent estimate (in the just-identified instrument case or our model).[4] This also permits improvements in the robustness of the estimate when the independent variables may suffer measurement errors (as noted in Appendix A). All variances were calculated using White's [1980] heteroscedasticity invariant-variance covariance matrix. Also, we tested for the presence of a unit root in these series, which were uniformly rejected, which we believe due to the relatively brief sample period. Results for the aggregate and state level sectoral estimates are provided in Tables 4, 5 and 6 in the text.

Our findings in these estimations also speak broadly to problems of endogeneity, or the direction of causation between incomes and broadband access that naturally plague studies of this type. While we have attempted to ameliorate this problem in our estimates through the technical application of the statistical technique of two-stage least squares, the strongest evidence that endogeneity is not an issue within our findings are in the results themselves. Since we find that broadband impacts are confined to two economic sectors (as opposed to impacts that are economy-wide), it is not inconsistent with the data to assume that broadband access within these sectors that is driving wages rather than the reverse.

Summary

The two econometric models offered in this section are derived from basic economic theory applied to the question of broadband. In the first model, we were able to clearly identify the role of broadband as it interacts with firm age to provide a productivity advantage. In the second model we were able to isolate impacts of broadband access on workers wages in two sectors: Services and FIRE, as suggested by

---

income and the number of residents with broadband access. We rejected COV(z,e) for z estimators at the .10 percent level.

[4] We employ population, number of residents with broadband access and per capita income as our instruments. See Green, Willim H 91997) Econometric Analysis, Prentice Hall. New York.

considerable earlier research. This finding, combined with the two stage least squares technique largely suggests that broadband is affecting wages in some sectors, and that our models do not suffer from the very real potential problem of reverse causation.

Contact:

Michael J. Hicks
hicksm@marshall.edu
Michael.hicks@afit.edu
Center for Business and Economic Research
Marshall University
1 John Marshall Drive
Huntington, WV 25755