Appendix A
Structural economic and econometric model

The Demand for Internet Access

The conventional labor-leisure choice model is extended to include the benefits from Internet access. The consumer is assumed to maximize a utility function of consumption and leisure, subject to a monetary budget constraint that includes the household production input *Internet bandwidth*, and subject to a time budget constraint that includes the household production input *time online*. Both inputs are used to produce reductions in essential time, defined as the non-remunerated time lost when participating in the labor market, plus time doing fundamental living activities such as banking, bill-paying, maintaining health, shopping, etc.

Essential time is represented by the household production function $\overline{T}(h, b, t; a)$, where h is the number of hours worked, b is Internet bandwidth, t is time spent online, and a is an efficiency parameter that reflects the technical ability of the individual. The function $\overline{T}$ is convex in b and t, and b and t are assumed to be complements in production so that increasing b will raise the marginal productivity of t. Similarly, a augments the productivity of b and t, decreasing essential time for a given input level. As such, $\overline{T}_b$, $\overline{T}_t$, $\overline{T}_a$, $\overline{T}_{bt}, \overline{T}_{ba}, \overline{T}_{ta} < 0$ and $\overline{T}_{bb}, \overline{T}_{tt} > 0$, where subscripts indicate partial derivatives. Some of the time costs of work may be fixed. Others, including commuting time, costs associated with the stress of work, the preparation and recovery period, and training and child care costs, may be linear or concave functions of the number of hours worked (Heim and Meyer, 2004). Essential time is concave in h so that $\overline{T}_h > 0$ and $\overline{T}_{hh} < 0$.

The consumer's maximization problem is:

$$\begin{aligned} \max_{h,b,t} \quad & U(c,\, L) \\ s.t. \quad & c = y + wh - p_b b - p_t t \\ & L = T - h - t - \overline{T}(h, b, t; a) \end{aligned} \tag{A1}$$

where U is utility, c is consumption, L is leisure, y is non-wage income, w is the wage rate, p_b is the per-unit price of bandwidth, p_t is the per-unit price of time online, and T is total time available.

Structural Econometric Models and Likelihoods

The individual's utility of an Internet service is assumed to be a function of the attributes of the service and a random error (known to the individual but not the researcher). This is the *Random Utility Model* (RUM) as it is applied in environmental economics, transportation research, health economics, and marketing.

It is assumed that respondents maximize their household's conditional utility of the service option (conditional on all other consumption and time allocation decisions):

$$U_{ij}^{k_{ij}} = \boldsymbol{\beta}'\boldsymbol{x}_{ij}^{k_{ij}} + \epsilon_{ij}^{k_{ij}} \;\; , \, i = 1, \ldots, n; \; j = 1, \ldots J, \; k_{ij} = 1, 2 \tag{A2}$$

where $U_{ij}^{k_{ij}}$ is the utility of alternative k_{ij} chosen by individual i during occasion j.[1] The vector $\boldsymbol{x}_{ij}$ contains the observed attributes of the alternatives. It is assumed that the $\epsilon_{ij}^{k_{ij}}$ are independent, and identically distributed mean zero normal random variables, uncorrelated with $\boldsymbol{x}_{ij}$, with constant unknown variance σ_{ϵ}^2.[2] The probability of choosing alternative 1, for example, is:

$$\begin{aligned} P_{ij}^1 &= P(U^1 > U^2) \\ &= P(\boldsymbol{\beta}'\boldsymbol{x}_{ij}^1 + \epsilon_{ij}^1 > \boldsymbol{\beta}'\boldsymbol{x}_{ij}^2 + \epsilon_{ij}^2) \\ &= P\Big(\epsilon_{ij}^2 - \epsilon_{ij}^1 < \, -\boldsymbol{\beta}'(\boldsymbol{x}_{ij}^2 - \boldsymbol{x}_{ij}^1)\Big) \\ &= \Phi\Big[-\boldsymbol{\beta}'(\boldsymbol{x}_{ij}^2 - \boldsymbol{x}_{ij}^1)/\sqrt{2}\sigma_{\epsilon}\Big] \end{aligned} \tag{A3}$$

and similarly for alternative 2, where $\sqrt{2}\sigma_{\epsilon}$ is the standard deviation of $\epsilon_{ij}^2 - \epsilon_{ij}^1$ and $\Phi(\,\cdot\,)$ is the univariate standard normal cumulative distribution function. Note that equation A2 comprises the usual probit model for dichotomous choice under the assumption the individual knows the random component and maximizes utility. The parameter vector $\boldsymbol{\beta}$ is identified only up to the scale factor $\sqrt{2}\sigma_{\epsilon}$, and σ_{ϵ} is not identified, since only the sign and not the scale of the dependent variable (the utility difference) is observed. If the J observations for each respondent are simply "stacked" to produce a data set with Jn observations, the unit of observation is an i, j pair and the likelihood is the product of the Jn probabilities like equation A2:

$$L(k_{ij}, \, i = 1, \ldots, n, \, j = 1, \ldots, J \,|\, \boldsymbol{x}_{ij}^1, \boldsymbol{x}_{ij}^2; \boldsymbol{\beta}, \sigma_{\epsilon}) = \prod_{i=1}^{n}\prod_{j=1}^{J} P_{ij}^{k_{ij}}. \tag{A4}$$

Incorporating the Status Quo Question

After choosing k_{ij}, individuals answer a question stating whether alternative k_{ij} would be chosen over the status quo. Let the status quo be indicated by 0. There are now four kinds of observations. Let the binary variable Z_{ij}^1 indicate the choice of alternative 1 or 2

---

[1]This notation, especially the use of k_{ij} to indicate either a 1 or a 2, is a bit cumbersome at first, but will make precise many of the concepts below.

[2]We allow for correlation of errors for an individual when it comes to choices involving the status quo–see section 3.2. For the hypothetical choices, there is no question of correlation since the effective errors that enter the likelihood are the *difference* in the two errors for any choice occasion, and the attribute sets are randomly assigned to choice "A" or choice "B". That is, the relevant distribution theory for forming the likelihood is based on $\epsilon_{i1}^1 - \epsilon_{i1}^2$, for example (person i, first choice occasion–see equation A7). In addition, any additive systematic component of the error is then eliminated. This is similar to the arguments of Heckman and Robb (1985) in their evaluation of social interventions.

for individual i on occasion j, and let the binary variable Z^2_{ij} indicate the chosen alternative or the status quo. These are defined by:

$$Z^1_{ij} = \begin{cases} 0 & \text{choose 1} \\ 1 & \text{choose 2} \end{cases} \qquad Z^2_{ij} = \begin{cases} 0 & \text{choose 1 or 2 over status quo} \\ 1 & \text{choose status quo over 1 or 2} \end{cases} \tag{A5}$$

Note that there is an information asymmetry here: when the status quo is chosen over 1 or 2 ($Z^2_{ij} = 1$), a complete ranking of the three alternatives has been determined; when 1 or 2 is chosen over the status quo ($Z^2_{ij} = 0$), all that is known is that 1 or 2 is the most preferred alternative.

Utility for the status quo, U^0_i under the model assumption (equation A1) is given by:

$$U^0_i = \boldsymbol{\beta}'\boldsymbol{x}^0_i + \epsilon^0_i \ , \tag{A6}$$

where ϵ^0_i are disturbances and $\boldsymbol{x}^0$ are the attributes of the individual's current Internet access. The attributes of the status quo vary over individuals, but not over choice occasions, and the utility of the status quo is evaluated only once by each individual (U^0_i and ϵ^0_i are subscripted with i only). The ϵ^0_i are assumed to be independent, identically distributed normal random variables with zero expectation and variance σ^2_0, uncorrelated with $\epsilon^{k_{ij}}_{ij}$.

The probability of choosing alternative k_{ij} $(1, 2)$ over alternative $3 - k_{ij}$ $(2, 1)$ and then choosing alternative k_{ij} over the status quo ($Z^2_{ij} = 0$) is the bivariate probability:

$$\begin{aligned} & P(U^{k_{ij}}_{ij} > U^{3-k_{ij}}_{ij}, U^{k_{ij}}_{ij} > U^0_i) \\ & = P\Big(\epsilon^{3-k_{ij}}_{ij} - \epsilon^{k_{ij}}_{ij} < -\boldsymbol{\beta}'(\boldsymbol{x}^{3-k_{ij}}_{ij} - \boldsymbol{x}^{k_{ij}}_{ij}), \epsilon^0_i - \epsilon^{k_{ij}}_{ij} < -\boldsymbol{\beta}'(\boldsymbol{x}^0 - \boldsymbol{x}^{k_{ij}}_{ij})\Big) \\ & = \Phi_2\Big[-\boldsymbol{\beta}'(\boldsymbol{x}^{3-k_{ij}}_{ij} - \boldsymbol{x}^{k_{ij}}_{ij})/\sqrt{2}\sigma_\epsilon, \ -\boldsymbol{\beta}'(\boldsymbol{x}^0 - \boldsymbol{x}^{k_{ij}}_{ij})/\sqrt{\sigma^2_0 + \sigma^2_\epsilon}; \ \rho\Big] \end{aligned} \tag{A7}$$

where ρ is the correlation between $\epsilon^{3-k_{ij}}_{ij} - \epsilon^{k_{ij}}_{ij}$ and $\epsilon^0_i - \epsilon^{k_{ij}}_{ij}$,

$$\rho = \frac{\sigma^2_\epsilon}{\sqrt{2\sigma^2_\epsilon(\sigma^2_0 + \sigma^2_\epsilon)}} = \frac{\sigma_\epsilon}{\sqrt{2(\sigma^2_0 + \sigma^2_\epsilon)}}, \tag{A8}$$

and Φ_2 is the standard bivariate normal cumulative distribution function. Similarly, the probability of choosing alternative k_{ij} over alternative $3 - k_{ij}$ and then choosing the status quo over alternative k_{ij} ($Z^2_{ij} = 1$) is:

$$\begin{aligned} & P(U^{k_{ij}}_{ij} > U^{3-k_{ij}}_{ij}, U^{k_{ij}}_{ij} < U^0_i) \\ & = P\Big(\epsilon^{3-k_{ij}}_{ij} - \epsilon^{k_{ij}}_{ij} < -\boldsymbol{\beta}'(\boldsymbol{x}^{3-k_{ij}}_{ij} - \boldsymbol{x}^{k_{ij}}_{ij}), \epsilon^0_i - \epsilon^{k_{ij}}_{ij} > -\boldsymbol{\beta}'(\boldsymbol{x}^0 - \boldsymbol{x}^{k_{ij}}_{ij})\Big) \\ & = \Phi_2\Big[-\boldsymbol{\beta}'(\boldsymbol{x}^{3-k_{ij}}_{ij} - \boldsymbol{x}^{k_{ij}}_{ij})/\sqrt{2}\sigma_\epsilon, \boldsymbol{\beta}'(\boldsymbol{x}^0 - \boldsymbol{x}^{k_{ij}}_{ij})/\sqrt{\sigma^2_0 + \sigma^2_\epsilon}; \ -\rho\Big] \end{aligned} \tag{A9}$$

where the symmetry of the normal distribution has been utilized.

One normalization is required: let $\sigma_\epsilon = 1/\sqrt{2}$. Define $\lambda^2 = \sigma_0^2/\sigma_\epsilon^2 = 2\sigma_0^2$. Then equation A8 can be written as:

$$P(U_{ij}^{k_{ij}} > U_{ij}^{3-k_{ij}}, U_{ij}^{k_{ij}} < U_i^0) = \tag{A8$'$}$$
$$\Phi_2\Big[-\boldsymbol{\beta}'(\boldsymbol{x}_{ij}^{3-k_{ij}} - \boldsymbol{x}_{ij}^{k_{ij}}), \frac{\boldsymbol{\beta}'(\boldsymbol{x}^0 - \boldsymbol{x}_{ij}^{k_{ij}})}{\sqrt{(1+\lambda^2)/2}}; -\frac{1}{\sqrt{2(1+\lambda^2)}}\Big]$$

and similarly for equation A6. The additional parameter to be estimated is λ. When $\lambda = 1$, $\sigma_\epsilon^2 = \sigma_0^2$ and the A versus B question and the question comparing A or B to the status quo have equal weight in the likelihood. When $\lambda < 1$ the question relating to the status quo contains more information, as there is more variability in the errors for the A *vs.* B question ($\sigma_\epsilon^2 > \sigma_0^2$), and conversely. Let $\boldsymbol{x}_{ij}^{rp} = (\boldsymbol{x}_{ij}^r - \boldsymbol{x}_{ij}^p)$ for $r,\ p = 0, 1$. Then the probabilities of the four data types are:

$$P(Z_{ij}^1 = 0, Z_{ij}^2 = 0) = \Phi_2\Big[-\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{21}, -\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{01}/\lambda; \frac{1}{2\lambda}\Big] \tag{A10}$$
$$P(Z_{ij}^1 = 0, Z_{ij}^2 = 1) = \Phi_2\Big[-\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{21}, \boldsymbol{\beta}'\boldsymbol{x}_{ij}^{01}/\lambda; -\frac{1}{2\lambda}\Big]$$
$$P(Z_{ij}^1 = 1, Z_{ij}^2 = 0) = \Phi_2\Big[\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{21}, -\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{02}/\lambda; \frac{1}{2\lambda}\Big]$$
$$P(Z_{ij}^1 = 1, Z_{ij}^2 = 1) = \Phi_2\Big[\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{21}, \boldsymbol{\beta}'\boldsymbol{x}_{ij}^{02}/\lambda; -\frac{1}{2\lambda}\Big]$$

The likelihood is the product of these Jn probabilities:

$$L(Z_{ij}^1, Z_{ij}^2,\ i = 1, \ldots, n,\ j = 1, \ldots, J|\boldsymbol{x}_{ij}^1, \boldsymbol{x}_{ij}^2, \boldsymbol{x}^0; \boldsymbol{\beta}, \lambda) = \prod_{i=1}^{n}\prod_{j=1}^{J} P(Z_{ij}^1, Z_{ij}^2) \tag{A11}$$

which, upon substitution of equations 9 can be written

$$L(Z_{ij}^1, Z_{ij}^2,\ i = 1, \ldots, n,\ j = 1, \ldots, J|\boldsymbol{x}_{ij}^1, \boldsymbol{x}_{ij}^2, \boldsymbol{x}^0; \boldsymbol{\beta}, \lambda) = \tag{A12}$$
$$\prod_{i=1}^{n}\prod_{j=1}^{J}\Phi_2\Big\{(-1)^{1-Z_{ij}^1}\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{21}, (-1)^{1-Z_{ij}^2}\Big[(1-Z_{ij}^1)\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{01} + Z_{ij}^1\boldsymbol{\beta}'\boldsymbol{x}_{ij}^{02}\Big]/\lambda; (-1)^{Z_{ij}^2}\frac{1}{\lambda}\Big\}$$

References

Heckman J. J., and R. Robb (1985). Alternative methods for evaluating the impact of interventions. In Longitudinal Analysis of Labor Market Data (eds. J. J. Heckman and B. Singer), 156-245. Cambridge: Cambridge University Press.

Heim, B. and Meyer, B. (2004), “Work Costs and Nonconvex Preferences in the Estimation of Labor Supply Models,” *Journal of Public Economics*, 88, 2323-2338.

Appendix B
Estimating the standard error of WTP measures from discrete choice experiments

Ignoring interactions, the utility model for Internet access choice is

$$U^*_{ij} = \beta_p p_{ij} + \boldsymbol{X}'_{ij}\boldsymbol{\beta}_a + \beta_s b_{ij} + \epsilon_{ij},\ i = 1, \ldots, n;\ j = 1, \ldots, 8. \tag{B1}$$

where p_{ij} is price, b_{ij} is bandwidth, and $\boldsymbol{\beta}_a$ is a $K \times 1$ vector of attributes of the service other than price and bandwidth. The estimates of WTP for these attributes are $\widehat{\boldsymbol{\beta}}_a/\widehat{\beta}_p$ and the estimated WTP for bandwidth is $\widehat{w}_b = \widehat{\beta}_s/\widehat{\beta}_p$.

Since the estimates of willingness-to-pay are nonlinear function of parameter estimates, their exact standard errors are unknown. While it would be possible to bootstrap the distribution of these estimators, since the normally distributed estimator of β_p is the denominator, the simulation would not converge to anything useful (see Kling and Sexton, 1990; Morey and Waldman, 1994). Instead, we use a linear approximation to the variance (sometimes known as the "delta method"). This approximation for elasticities has been examined in Krinsky and Robb (1986).

Define the $(K+1) \times 1$ vector

$$\widehat{\boldsymbol{w}} = (\ \widehat{\boldsymbol{\beta}}_a \vdots\ \widehat{\beta}_s)/\widehat{\beta}_p\ \ . \tag{B2}$$

Define the $(K+2) \times 1$ vector of parameter estimates $\widehat{\boldsymbol{\theta}} = \left(\widehat{\beta}_p \vdots \widehat{\boldsymbol{\beta}}'_a \vdots \widehat{\beta}_s\right)'$. Let $\widehat{\boldsymbol{\Sigma}}$ be the estimated variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$. The linear approximation to the variance of $\widehat{\boldsymbol{w}}$ is

$$\widehat{V}(\widehat{\boldsymbol{w}}) \approx \left[\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\theta}}\right]' \widehat{\boldsymbol{\Sigma}} \left[\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\theta}}\right] \tag{B3}$$

where the derivatives are evaluated at the parameter estimates. The square root of the diagonal elements of $\widehat{V}(\widehat{\boldsymbol{w}})$ are the estimated standard errors of the estimates of WTP. These derivatives are

$$\frac{\partial \widehat{\boldsymbol{w}}}{\partial \widehat{\boldsymbol{\theta}}} = \begin{pmatrix} -\frac{\widehat{\beta}_{a_1}}{\widehat{\beta}_p^2} & \frac{1}{\widehat{\beta}_p} & 0 & 0 & \ldots & 0 \\ -\frac{\widehat{\beta}_{a_2}}{\widehat{\beta}_p^2} & 0 & \frac{1}{\widehat{\beta}_p} & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ -\frac{\widehat{\beta}_{a_K}}{\widehat{\beta}_p^2} & 0 & 0 & 0 & \ldots & 0 \\ -\frac{\widehat{\beta}_s}{\widehat{\beta}_p^2} & 0 & 0 & 0 & \ldots & \frac{1}{\widehat{\beta}_p} \end{pmatrix} \tag{B4}$$

Focusing on bandwidth, the estimated variance of the WTP for bandwidth from equation B2 is

$$V(\widehat{w}_s) = \Big(\frac{\widehat{\beta}_s^2}{\widehat{\beta}_p^4}\Big)\widehat{\sigma}_{pp} + 2\Big(\frac{\widehat{\beta}_s}{\widehat{\beta}_p^3}\Big)\widehat{\sigma}_{ps} + \frac{1}{\widehat{\beta}_p^2}\widehat{\sigma}_{ss}$$

The utility model for access, with interactions, is

$$U_{ij}^* = \beta_p p_{ij} + \boldsymbol{X}_{ij}'\boldsymbol{\beta} + (\beta_s + \boldsymbol{a_i'\delta})b_{ij} + \epsilon_{ij},\ i = 1,\ldots,n;\ j = 1,\ldots,8, \qquad \text{(B5)}$$

where $\boldsymbol{a}_i$ is a vector of L demographic variables for individual i and the elements of $\boldsymbol{\delta}$ are additional parameters to be estimated. The estimate of WTP for bandwidth from this model is

$$\widehat{w}_s = \ (\widehat{\beta}_s + \boldsymbol{\overline{a}_i'}\widehat{\boldsymbol{\delta}})/\widehat{\beta}_p \qquad \text{(B6)}$$

where the vector of individual-specific demographic variables is evaluated at their means. Define $\widehat{\boldsymbol{\phi}} = \left(\widehat{\beta}_p \vdots \widehat{\beta}_s \vdots \widehat{\boldsymbol{\delta}}'\right)'$, and define $V\left(\widehat{\boldsymbol{\phi}}\right) = \boldsymbol{\Sigma}^*$. The variance of $\widehat{w}_s$ is

$$\widehat{V}(\widehat{w}_s) \approx \left[\frac{\partial \widehat{w}_s}{\partial \boldsymbol{\phi}}\right]' \widehat{\boldsymbol{\Sigma}}^* \left[\frac{\partial \widehat{w}_s}{\partial \boldsymbol{\phi}}\right] \qquad \text{(B7)}$$

where

$$\frac{\partial \widehat{w}_s}{\partial \boldsymbol{\phi}} = \begin{pmatrix} -\frac{\widehat{\beta}_s + \boldsymbol{\overline{a}_i'}\widehat{\boldsymbol{\delta}}}{\widehat{\beta}_p^2} & \frac{1}{\widehat{\beta}_p} & \frac{a_1}{\widehat{\beta}_p} & \frac{a_2}{\widehat{\beta}_p} & \cdots & \frac{a_L}{\widehat{\beta}_p} \end{pmatrix}' . \qquad \text{(B8)}$$

Reference:

Kling, C., and R. Sexton (1990). “Bootstrapping in Applied Welfare Analysis.” *American Journal of Agricultural Economics* 72: p.

Krinsky, I., and A. Robb (1986). “On Approximating the Statistical Properties of Elasticities.” *Review of Economics and Statistics* 68(4): p. 715-19.

Morey, E., and D. Waldman (1994). “Functional Form and the Statistical Properties of Welfare Measures–A Comment.” *American Journal of Agricultural Economics* 76(4): p. 954-57.

Appendix C
Details on the study design: within subjects

The likelihood as it is written in equation A12 does not take into consideration the fact that the formation of that part of the likelihood involving the comparison of the chosen alternative to the status quo involves the error difference $\epsilon_i^0 - \epsilon_{ij}^{k_{ij}}$, where $k_{ij} = 1$ or 2 (depending upon the choice), and from choice occasion to choice occasion these error differences are correlated. This correlation is induced by the common occurrence of ϵ_i^0, since respondents need evaluate their utility of the status quo only once. This point is generally missed in conjoint analysis. An econometric innovation of this study is to treat the person, and not the person-choice occasion, as the unit of observation, so that we may explicitly model this correlation. The likelihood is now written

$$L(Z_{ij}^1, Z_{ij}^2,\, i=1,\ldots,n,\, j=1,\ldots,J|\boldsymbol{x}_{ij}^1, \boldsymbol{x}_{ij}^2, \boldsymbol{x}^0; \boldsymbol{\beta}, \lambda) = \prod_{i=1}^{n} P(Z_{i1}^1, Z_{i1}^2, Z_{i2}^1, Z_{i2}^2, \ldots, Z_{iJ}^1, Z_{iJ}^2) \ . \tag{C1}$$

The probability in equation C1 would appear to be computationally intractable, as it involves a 16-fold $(2 \times J = 8)$ integration of the multivariate normal density function. Fortunately, this is not the case, as the correlation between $\epsilon_i^0 - \epsilon_{ij}^1$ and $\epsilon_i^0 - \epsilon_{ij}^2$, for example, is a result of the common occurrence of ϵ_i^0. This means that we can follow a familiar conditioning argument to express the probability in equation C1 as the integral of the product of eight bivariate probabilities, integrated against the univariate normal density (see Waldman, 1985). But the cost of this generality is in programming and computer time, as the likelihood must be maximized by simulation or with quadrature methods. We used Hermite polynomial quadrature (Abramowitz and Stegun, 1964, p. 890).

References

Abramowich, M., and Stegun, J. (1964). "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables." National Bureau of Standards, Applied Mathematics Series - 55.

Waldman, Donald M., 1985. "Computation in Duration Models with Heterogeneity." Journal of Econometrics, Vol. 28, 127-134.