APPENDIX TO BE PUBLISHED ONLINE

A. Monopoly assumption and FOCs

Monopoly assumption: The operator whose data we use had a considerable degree of monopoly power over the first (actual and potential) users of the mobile internet services.[1] The primary reason for this is that most of the potential users were captive. There are four pieces of evidence supporting this claim: First, the consumers had subscribed to a wireless-voice plan before the new mobile services were available. This means that once available, the new service was a kind of add-on and consumed within the plan, i.e., conditional on being a subscriber to the particular voice-dominated service bundle. As we will argue, the pricing of the add-on (i.e., the new service within the bundle) was at the discretion of the operator. Second, over the period our data are from, the voice and earlier non-voice (e.g., text messaging) parts of the bill were clearly larger than that emanating from the use of mobile internet services. We therefore think it unlikely that a different price for the new mobile Internet services would have lead customers to switch operators and by implication, that there was any significant price competition in the dimension of the operator's offering we focus on. Third, even if a consumer wanted to switch, there was a considerable non-pecuniary hurdle: The cellular phone numbers were not portable at the time. According to survey evidence reported by the Finnish Ministry of Transports and Communications, portability was the most important obstacle to switching (see also Viard 2006). The statistics on the number of consumers who have switched after portability was implemented in June 2003 unambiguously support this survey evidence. Finally, assuming monopoly pricing will give us a lower bound estimate of marginal costs. Our prior is that prices were too high, implying that monopoly assumption works against our prior. The results confirm ex post that the monopoly assumption is reasonable in the sense of us finding that the actual prices were higher than static optimal monopoly prices. That is, the estimated optimal *static monopoly* prices from our preferred model are *lower* than the actual prices that were charged. This

[1] To be sure, it was not strictly speaking a formal monopoly. The operator was over the time period we study one of the two dominant players in the Finnish mobile telecommunications market.

revealed pricing behavior strongly suggests that the operator whose data we have did *not* behave as if it faced significant price competition over the users of the early mobile services.[2]

Optimal two-part tariff and first-order conditions (FOCs): Ignoring consumer heterogeneity for brevity, the optimal two-part tariff solves

$$\text{(A.1)} \qquad \max_{p,K} \Omega = \left[\mu(1-F(\bar{\alpha}(\tau p, \tau K)))\right]\left[q_{\alpha>\bar{\alpha}}(\tau p, \tau K)(p-c)+K-C\right]$$

where c = marginal cost per minute of connection, C = marginal cost of opening/terminating a connection, $q_{\alpha>\bar{\alpha}}$ = conditional connection length, and τ = the valued added tax (VAT) parameter (as VAT is 22%, τ = 1.22), implying that per-minute price (connection fee) charged from the customer is τp (τK). The above objective function also explicitly allows for the dependence of $\bar{\alpha}$ and $q_{\alpha>\bar{\alpha}}$ on the parameters of the tariff.

Let

$A(p,K) \equiv \mu(1-F(\bar{\alpha}(\tau p, \tau K)))$; and

$B(p,K) \equiv q_{\alpha>\bar{\alpha}}(\tau p, \tau K)(p-c)+K-C$.

The general form of the FOCs is

$$\text{(A.2)} \qquad \frac{\partial \Omega}{\partial i} = \frac{\partial A(p,K)}{\partial i} B(p,K) + A(p,K)\frac{\partial B(p,K)}{\partial i} = 0$$

for i = *p, K*. For the linear demand function, the parts of the FOCs can be written as:

$$\text{(A.3)} \qquad A(p,K) = \mu \exp(-\lambda(\gamma\tau p + \sqrt{2\gamma\tau K})$$

$$\text{(A.4)} \qquad B(p,K) = \left[\sqrt{2\gamma\tau K} + \frac{1}{\lambda}\right](p-c) + K - C$$

$$\text{(A.5)} \qquad \frac{\partial A(p,K)}{\partial p} = -\lambda\gamma\tau\mu(1-F(\bar{\alpha}(\tau p, \tau K)))$$

[2] If and when rivals offer similar products at different prices, changing (raising) prices might lead some customers to change their operator. The estimated optimal monopoly prices mean, however, that should our operator have a lower degree of monopoly power than we have assumed, a change towards the optimal prices should have induced an in-flow of customers (assuming the potential rivals would not have changed their behavior). If rivals had matched the price changes, there would have been no reason for their customers to change operators. It is therefore unlikely that our customer-level figures would be greatly affected by such switches. If at all, the effect would most likely be that our figures present a lower bound. The reason for this is that the most likely customers to change from a rival operator to the operator whose data we have would have been those with above average valuation (and hence usage) of mobile internet services.

(A.6) $$\frac{\partial B(p,K)}{\partial p} = \sqrt{2\gamma\tau(1/K)} + 1/\lambda$$

(A.7) $$\frac{\partial A(p,K)}{\partial K} = \left[-(1/2)\sqrt{2\gamma\tau(1/K)}\right](1 - F(\bar{\alpha}(\tau p, \tau K)))$$

(A.8) $$\frac{\partial B(p,K)}{\partial K} = (1/2)\sqrt{2\gamma\tau K}(p-c) + 1 .$$

For the log- linear demand function, the corresponding parts take the following forms:

(A.9) $$A(p,K) = \mu \exp(-\lambda\gamma\tau^2 K \exp(\gamma\tau p))$$

(A.10) $$B(p,K) = [(p-c)(\gamma K\tau^2 \exp(\gamma\tau p) + 1/\lambda)\exp(-\gamma\tau p) + K - C]$$

(A.11) $$\frac{\partial A(p,K)}{\partial p} = -\lambda\gamma^2\tau^3 K \exp(\gamma\tau p)\left[\mu(1 - F(\bar{\alpha}(\tau p, \tau K)))\right]$$

(A.12) $$\frac{\partial B(p,K)}{\partial p} = \gamma K\tau^2 + \exp(-\gamma\tau p)(1/\lambda)(1 - \gamma\tau(p-c)]$$

(A.13) $$\frac{\partial A(p,K)}{\partial K} = -\lambda\gamma\tau^2 \exp(\gamma\tau p)\left[\mu(1 - F(\bar{\alpha}(\tau p, \tau K)))\right]$$

(A.14) $$\frac{\partial B(p,K)}{\partial K} = \gamma\tau^2(p-c) + 1 .$$

B. Elicitation of expert opinions

Elicitation procedure: In initial discussions with an industry expert it became clear that the relevant marginal costs are, given the circumstances of WAP introduction (i.e., existing networks with high coverage, large capacity and therefore no need to install new capacity/expansion of network coverage), the opening/termination cost of an extra WAP (GSM) connection through the existing network, and the marginal cost of lengthening an existing WAP (GSM) connection by one more minute. It also turned out that it is most likely the case that the operators had not systematically gathered and/or stored data on the two marginal costs.

To obtain information on marginal costs, we designed a structured survey that consists in total of 13 questions which we personally administrated to a number of acknowledged industry experts (more on this below). The structure of the elicitation procedure is as follows: First, a definition and an example of marginal costs is given and discussed, if needed, at length with the industry expert. Second, the respondents are asked for point estimates of the two marginal costs (Q1-Q2). The next two questions are about the marginal costs of GSM calls (Q3-Q4) and the fifth (Q5) was a statement about the (technical) equivalence of WAP connections and GSM calls. This question (Q5) is a check-question, as it allows us to separate the answers of those respondents who know that WAP connections were technically identical to GSM calls from the answers of those who do not know this. The sixth question (Q6) is about the elasticity of demand and the seventh (Q7) about pricing. Finally, a series of questions (Q8-Q13) elicits information about the distribution of marginal costs, following the work of Dominitz and Manski (1996, 1997). This elicitation procedure was pre-tested with an industry expert and modified slightly on the basis of his reactions to it.

Selection of the industry experts: We asked for names of people likely to have intimate knowledge of WAP as of 2001 both from an acknowledged industry expert, and from operators and a large telecom equipment company, yielding us in the end 36 names. We sent these people an email explaining our survey (with the survey attached and the elicitation procedure described) and requested to call them at their convenience. After a couple of reminder emails we phoned them if we had not received any reaction. When making the call, we explained them (as described in the survey document) the purpose of the study and the survey, and proceeded to ask the questions. We did not systematically record the length of the interviews, but a typical interview lasted 35-45 minutes.

Response rate and reasons for non-response: In the end, 18 people agreed to answer our questions, but several of them made various reservations. Those who declined invariably gave as the main reason that they felt that they did not have adequate expertise to answer our questions (which they often called difficult). Only one person declined because of corporate policy. We ended up discarding answers of 5 people because they did not agree at least partly with our technical background question (Q5 in the survey). Some of those who are included in the final survey sample made various reservations which we recorded. In order to be conservative, we decided to keep their answers in the sample as they regularly gave higher estimates of marginal costs than those respondents not making reservations. By doing this we hope to have avoided any downward bias in our estimated marginal costs.

C. Estimation results of the baseline model

We present the basic estimation results of the first and second step of the two-step m-estimation procedure briefly here, because the key insights come from the economic implications of the model. Table A.1 displays the cross-validation results, separately for $\exp(g(\underset{\sim}{z}_i, \underset{\sim}{\mu}))$ and $k(\underset{\sim}{z}_i, \underset{\sim}{\lambda})$.

The first step estimation results are displayed in Table A.2, which among other things shows that Wald-tests indicate that the included variables are jointly highly significant. The results from the second step are displayed in Table A.3, separately for linear and log-linear demand. We report Huber-White standard errors adjusted for clustering within consumers in Table A.2 and bootstrapped standard errors in Table A.3. Bootstrapping the standard errors in the second step is important, for they are about ten times larger than the unadjusted (incorrect) standard errors.

[Insert Tables A.1, A.2 and A.3 here]

D. Comparison to i-mode

A comparison of the success of the early WAP with the early diffusion of i-mode, a service brand of NTT DoCoMo, which took off quite rapidly after its introduction in February 1999 in Japan, provides a reality check on whether quality-adjusted WAP prices were high.

First, i-mode's overall usability appears to have been somewhat better than that of WAP. It was, for example, a bit faster, and supported colors. Second, when i-mode was introduced, the take up of wireline Internet was relatively modest in Japan compared to Western Europe, making its content relatively novel. Finally, it seems that a number of the services made available by i-mode were *not* widely available in Japan before, suggesting less repackaging.[3] All this suggests that i-mode had fewer substitutes and better quality than WAP.

For a back-of-the-envelope i-mode calculation, let us start by emphasizing that reliable price comparisons between WAP and i-mode are difficult due to differences in the pricing principles (i-mode's package based vs. WAP's two-part metered tariff) and especially due to lack of comparable data: From Kakkori (2001) we can infer that in Spring 2001, the monthly fee for the i-mode service was 2.73 euros. In addition, one packet, i.e., 128 bytes, cost 0.0027 euros. If a representative WAP connection in our data had involved a transfer of 2 kilobytes, the price of making the connection using i-mode would have been around 0.042 euros. Comparing the i-mode price to the prices used and obtained in our counterfactual calculations show that usage would have been higher than it actually was. If one took into account differences in service quality (see the main text), the increase in usage would correspondingly be larger.

[3] Examples of new services are messages across operators' networks, and picture downloads. Especially entertainment related services, such as "What's new" -information services and music sites, became popular early on (Marketing Interactive Network, 2000).

E. Further robustness checks: Learning, consumption shifting and advertising

In Table A.4 we report reduced form estimations of Poisson models of connection counts, and within estimates of connection length, conditioning naturally on strictly positive connection length. Both specifications include consumer fixed effects and period fixed effects with period 1 as the base period. As can be seen from the table, the number of connections in one of the two post-experiment periods was smaller than in period 1 (the base period), and larger in the other. Although this doesn't strictly rule out experimentation, it does suggest that the consumers did not update their beliefs regarding the quality of their goods. This, coupled with the fact that the number of connections remains high through all three experiment periods suggests that learning or experimentation is not a culprit: If it were the case that consumers initially experimented, they should have rather quickly have realized that their priors hold, in which case the demand during the later parts of the experiment (period 5) should already mostly reflect "true" demand. The demands during the three experimental periods, while decreasing in time, are rather close to each other. While re-estimation of the model using only period 5 (i.e., excluding data from the first two experiment periods 3 and 4) data produces lower elasticities and higher consumer surpluses, the overall picture remains the same.

[Insert Table A.4 here]

As to shifting consumption into the experiment period, the results do not support that alternative either as the coefficients of both the last pre-experiment and the first post-experiment periods are positive, indicating higher usage than during the base period (period 1). Finally, as the post-experiment number and length of connections is not systematically higher than in the pre-experiment periods, there is no evidence that advertising resulted in a shift in demand.

We repeated this analysis using a sample of individuals who early on reacted strongly to the pricing experiment. The reasoning is that if some individuals only learned of the experiment in the 2^{nd} or 3^{rd} two-week period, they might not have had enough time to experiment/learn. We thus computed for each customer the change in number of connections from the last pre-experiment to the 1^{st} experiment period, and excluded all individuals who were below the mean (in the second variant, median) change. We then reran the reduced form fixed effects Poisson and length of call estimations including only the last experiment period and all the non-experiment periods in the estimation sample. The results, available upon request, are very similar to those reported above.

F. Results using data from plan B consumers

We repeat our empirical analysis using the data on the plan B subscribers. For brevity, we only note that cross-validation results suggest that the simplest model (Model 1 in Table A.2) suffices. The estimated latent demand remains strong and price elasticities high: The average number of needs per consumer is 14.36 per a two-week period. The average per-minute price (p_t) elasticity of connection length ranges from -2.04 (log-linear demand) to -2.91 (linear demand). We also obtain large price elasticities of the number of wireless connections for the per-minute price and the fixed fee. For example, the per-minute price elasticity of the number of wireless connections is on average -1.78, and ranges from -0.82 to -3.20 for the linear demand. However the conditional probability of opening a connection is (again) small, on average about 9%. When translated into annual numbers, the per consumer surplus is on average 18.33 euros in the case of linear demand and 1.54 euros in the case of log-linear demand. These results are in line with the results we obtained using the data on the more popular plan.

G. Reduced form (consumer surplus) estimates

In our case, the relevant consumer surplus or compensating variation (CV) -measure is the difference in the consumer's expenditure function between the expenditure at the market prices and at the service's virtual price, which is the price that sets the service's demand to zero. These expenditures are measured at the level of the utility received once the new service is on the market. Hausman (1999) shows that a practical way to compute the welfare gain is to use the approximation $CV \approx -0.5pY/\varepsilon$, where Y is the quantity consumed, p is the price, and ε is the own-price elasticity. In our case, direct application of the above approach is complicated for three reasons. First, what is Y? Second, what is p? Third, what functional form should we use to estimate ε? The standard log-log linear form is an option, but it cannot in our case be linearized conveniently by taking logarithms, because Y is frequently zero, as is the price during the experiment periods.

We proceed by assuming that Y is the number of connections made. To define p we compute the average of connection lengths over all consumer-period observations for which the length is positive, using data from the non-experiment periods only. We then take as the imputed price per connection the sum of the fixed connection fee plus the average outlay per connections, defined as the product of the per-minute charge and the average connection length. This imputed total price is 0.42 euros per connection. We then regress the number of connections on the imputed total price using a standard Poisson regression. Table A.5 reports the own-price elasticity and consumer surplus estimates derived from the reduced form regression that we report in the main body of the text. It also gives the underlying regression coefficients.

[Insert Table A.5. here]

With this benchmark reduced form estimate at hand, we perform a number of additional reduced-form estimations: First, we replicate the results of Table A.5 by estimating a standard linear model with ordinary least squares. The coefficient of the price variable is -115.70, which results in even lower consumer surplus. Second, we use a different imputed per connection price: If we simply take the "total" price to be the fixed connection fee, which is 0.09 euros, plus the per-minute price, which in the more popular plan is 0.12 euros, we obtain an elasticity estimate that is clearly lower, about one fifth of that presented in Table A.5. The estimated consumer surpluses remain, however, negligible at 1.31 euros per consumer per annum. Finally, we allow for consumer heterogeneity: Estimating a fixed-effects Poisson (Hausman, Hall, and Griliches 1984) reinforces the finding of highly elastic demand: At the imputed total price of 0.42 euros per connection, the own-price elasticity is -10.83. Using the demographic variables as an alternative way to control for the heterogeneity produces very similar results. Taken together, these alternative welfare calculations echo the findings reported in the main text.

Table A.1. Cross-validation results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
CONSTANT	X	X	X	X	X	X
AGE	X	X	X	X	X	X
GENDER	X	X	X	X	X	X
CITY	X	X	X	X	X	X
GENDER*AGE		X	X	X	X	X
CITY*AGE		X	X	X	X	X
GENDER*CITY		X	X	X	X	X
AGE^2			X	X	X	X
GENDER*AGE^2				X	X	X
CITY*AGE^2				X	X	X
AGE^3					X	X
GENDER*AGE^3						X
CITY*AGE^3						X
Cross-validation						
of exp(g(.))	2588.502	2585.69	2588.231	2590.941	2591.215	2601.392
of k(.)	110.1632	110.1178	110.0678	110.0267	109.8905	110.023

Table A.2. Estimation results from the first step (μ and λ)

	WAP_COUNT (μ)		CALL_DUR (λ)	
	Coefficient	Std. error*	Coefficient	Std. error*
AGE	0.01	0.01	-0.07	0.04
GENDER	0.11	0.29	-0.05	0.32
CITY	0.40	0.27	-0.61	0.31
GENDER*AGE	-3.8E-03	6.5E-03	-7.3E-03	1.6E-02
CITY*AGE	-0.02	0.01	0.03	0.02
GENDER*CITY	0.05	0.16	-0.03	0.08
AGE2	-	-	1.5E-03	7.9E-04
GENDER*AGE2	-	-	1.7E-04	2.0E-04
CITY*AGE2	-	-	-3.1E-04	2.2E-04
AGE3	-	-	-1.1E-05	5.4E-06
CONSTANT	2.13	0.28	2.81	0.52
Obs.	5460		4350	
Wald (joint significance)	36.91		67.61	
d.f.	6		10	
p-value	0.000		0.000	
Log-likelihood	-66434.57		-5555.11	

*Huber-White heteroscedasticity robust covariance matrix, adjusted for clustering within consumers

Table A.3. Estimation results from the second step (γ)

	Log-linear demand		Linear demand	
	Coefficient	Std. error	Coefficient	Std. error
AGE	-0.10	0.04	-1.64	0.59
GENDER	-3.95	1.74	-62.92	28.36
CITY	-0.95	2.12	-9.00	34.03
GENDER*AGE	0.09	0.04	1.46	0.66
CITY*AGE	-0.03	0.05	-0.40	0.67
GENDER*CITY	-0.30	1.26	-12.12	18.30
CONSTANT	22.28	1.72	155.77	25.51
Obs.	12740		12740	

Note: Standard errors are produced using a bootstrap with 500 repetitions.

Table A.4. Reduced form estimation of connection count and call length (cond. > 0)

	Count		Length	
Period-dummy	Coefficient	p-value	Coefficient	p-value
t = 2	0.28	< 0.01	0.14	0.59
t = 3	2.65	< 0.01	2.78	< 0.01
t = 4	2.34	< 0.01	2.65	< 0.01
t = 5	2.25	< 0.01	2.44	< 0.01
t = 6	0.17	0.00	-0.41	0.12
t = 7	-0.09	0.01	0.14	0.61
Consumer fixed-effects	YES		YES	
Tests of equality of period coefficients (p-values)				
2 vs. 6	< 0.01		0.03	
2 vs. 7	< 0.01		0.99	
6 vs. 7	< 0.01		0.03	

Table A.5. Reduced form Poisson estimation and consumer surplus

	Own-price elasticity	Annual per-consumer surplus (euros)
At the imputed price:	-10.84	0.56

	Price-variable	Constant
Coefficient	-25.89	10.95
Standard error	0.85	0.29