The Journal of Choice Modelling 16 (2015) 15–22

Contents lists available at ScienceDirect

The Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples

Kenneth Train

Department of Economics, University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 27 May 2015
Received in revised form
13 August 2015
Accepted 11 September 2015

ABSTRACT

The attributes that consumers anticipate or expect when choosing among alternatives (i.e., prior to consumption) often differ from the attributes that they actually experience when consuming their chosen alternative. This paper describes, and illustrates with several examples, the calculation of consumer surplus in this situation and the loss in consumer surplus due to the imperfect foreknowledge about attributes. The procedures are useful in many settings, such as the assessment of damages for false advertising and the analysis of informational policies.

1. Introduction

In many choice situations, the consumer evaluates alternatives on the basis of the his/her concept of what the attributes of the alternatives are, but then finds, after making the choice, that the attributes of the chosen alternative are different than anticipated. The disconnect between anticipated and experienced attributes can arise from many sources: consumers might not accurately evaluate products (e.g., in choosing a car, consumers might not accurately translate differences in mpg into differences in operating cost); producers might misrepresent or withhold information about their products' attributes, or not know the attributes themselves (such as an undiscovered defect); or the attributes might depend on factors that simply cannot be known at the time of choice (e.g., vehicles' operating costs depend on future gas prices, which are not known at the time of choosing among vehicles.)[1] The situation can be said to arise from consumers' imperfect foreknowledge, but without any connotation that the imperfection is due to an error by the consumer.

The question that this paper addresses is how to measure consumer surplus in these situations, and particularly, how to measure the loss in consumer surplus that arises from imperfect foreknowledge of alternatives' attributes. This measure of loss has different uses depending on the source of the discrepancy between anticipated and experienced attributes. If consumer error is the source, then the loss represents, equivalently, the welfare gain that can be attained by correcting the error. If the discrepancy is due to false advertising, then the loss constitutes the compensation that is needed for the firm to atone for the incorrect information. And if the situation is unavoidable e.g. when attributes cannot be known with certainty in advance, then the measure of loss can be used to evaluate policies that alleviate the impact of the uncertainty.

E-mail address: train@econ.berkeley.edu

[1] In regard to this last source, the discrepancy is a natural outcome within the standard theory of decision-making under uncertainty. Consumers choose to maximize expected utility based on a probability distribution of possible attribute levels for each alternative. However, after the choice is made, some realization from the distribution is revealed, which is inevitably different from the expectation.

http://dx.doi.org/10.1016/j.jocm.2015.09.003
1755-5345/

Numerous papers have examined the role of perceptions in consumers' decisions, from both a general perspective (e.g., Kahneman, 1994; Manski, 2004) and in specific contexts (e.g., Delavande, 2008; van der Klaauw and Kenneth, 2008; Jensen, 2010).[2] However, the question of how to calculate consumer surplus when anticipated and experienced attributes differ has hardly been touched. I have found two papers that describe welfare calculation for this situation: Allcott (2013) and Schmeiser (2014). These authors give formulas similar to the ones I describe below and should be considered the originators of the procedure. However, their derivations are in the context of specific topics and constitute small sections of their papers, such that the general relevance could easily be missed by readers. The current paper gives the formulas prominence, explaining them in greater detail and providing illustrative applications. As will be shown, the calculations are straightforward for standard and mixed logit models, employing the familiar log-sum term and an additional easy-to-calculate term.

In many situations, the consumer can learn about product attributes – by seeking information, from experience with the chosen alternatives in repeated choices, and through the experiences of other consumers who make other choices – and this learning will reduce and perhaps eliminate the difference between anticipated and experienced attributes. The current paper does not address how consumers' concepts about attributes change over time. Rather, the procedures in this paper are useable for calculating consumer surplus at any given point in the learning process, and to represent the loss due to the gap that remains at that point.

2. Derivation of welfare measures

A person faces a choice among a set of discrete alternatives indexed $j=1,\dots J$. The utility the person expects from alternative j is W_j, which I call "anticipated utility." The utility that the person actually obtains is U_j, called "actual utility."[3] The difference between the actual and anticipated utility is $d_j = U_j - W_j$, such that $d_j > 0$ if the alternative is actually better than the person thought and $d_j < 0$ if worse. Kahneman (1994), Allcott (2013), and others have used equivalent terms, such as "belief utility" and "experience utility."

The person chooses among the alternatives based on anticipated utility but obtains the actual utility from the chosen alternative. The alternative that the person chooses is j^* for which $W_{j*} > W_j \ \forall\, j \neq j^*$.[4] The utility that the person attains from the choice is U_{j*}.

The alternative that provides the highest actual utility is k^* s.t. $U_{k^*} > U_j \ \forall\, j \neq k^*$. The loss due to the person's imperfect foreknowledge is then $U_{j*} - U_{k*}$. An implication, which can have a large impact on welfare calculations, is that a person loses from imperfect foreknowledge only if the person would have chosen a different alternative under perfect foreknowledge. If the same choice would have been made, such that $k^* = j^*$, then $U_{j*} - U_{k*} = 0$.

Suppose $k^* \neq j^*$ and $d_{j*} < 0$, such that the consumer incurs a loss due to their imperfect foreknowledge of alternatives ($k^* \neq j^*$) and a loss relative to anticipated utility ($d_{j*} < 0$). As long as $d_{k^*} \leq 0$, the loss due to imperfect foreknowledge is smaller in magnitude than the loss relative to anticipated utility:

$$U_{j*} - U_{k*} = W_{j*} - W_{k*} + d_{j*} - d_{k*} > d_{j*} - d_{k*} \geq d_{j*}$$

such that, since both terms are negative, $|U_{j*} - U_{k*}| < |d_{j*}|$.[5]

Measurement of these terms can be readily operationalized with a standard logit specification. Let anticipated utility be

$$W_j = -\alpha c_j + \beta x_j + \varepsilon_j$$

where c_j is the anticipated cost, x_j is a vector of other anticipated attributes, and ε_j is random and assumed to be iid extreme value. The choice probabilities are the standard logit formula based on anticipated utility:

$$P_j = \frac{e^{-\alpha c_j + \beta x_j}}{\sum_k e^{-\alpha c_k + \beta x_k}}$$

Actual utility is d_j different than anticipated utility: $U_j = W_j + d_j$. We assume, for the formulas below, that each ε_j is independent of the d_j's, in the same way that they are independent of c_j and x_j for all j.[6] Average consumer surplus from this

[2] Analysis of perceptions clearly relates to the current situation but can be more general. In particular, perceptions can affect equally the utility that the person anticipates and experiences. For example, people who think that owning a Tesla will make them look good in other people's eyes obtain utility from that perception whether or not it is true. The perception creates a difference between anticipated and experienced attributes only if any of these people discover after buying the car that other people haven't changed their opinions of them.

[3] I use the term "actual utility" instead of "experienced utility" for linguistic convenience. Stated in terms of experience, U_j is the utility that the person would experience if the person were to choose alternative j.

[4] Through the discussion, the possibility of ties is ignored, or, more precisely, is given a zero probability of occurring.

[5] The difference between the types of losses can have a large impact on legal actions. Suppose a firm is charged with false advertising. If found guilty, the firm might be required to compensate for consumers' loss relative to the utility they would have obtained without the false advertising, i.e., for $U_{j^*} - U_{k^*}$. If instead the firm is found guilty of a contract violation because of not providing the product that was promised, the firm might be required (depending on contract law) to compensate consumers for the loss relative to the promised product, which is d_{j*}.

[6] This assumption implies, importantly, that the distribution of the portion of anticipated utility that is unobserved by the researcher is the same as for actual utility. The errors for any one person need not be the same, but the distribution of errors over people is the same. An useful area for future research is

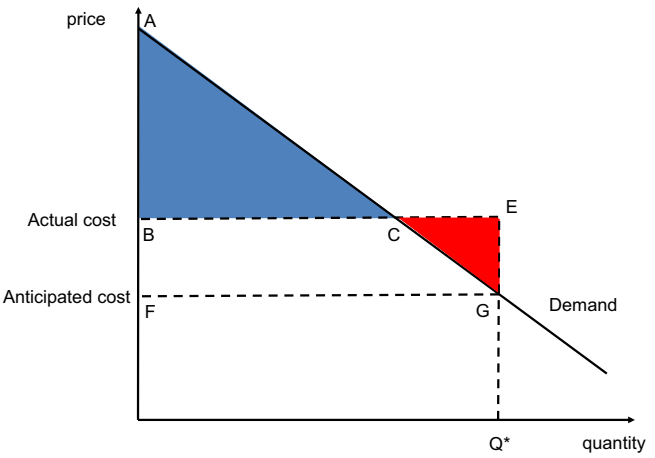

Fig. 1. Consumer surplus under imperfect foreknowledge about cost.

choice (averaged over a population of people with the same values, for each alternative, of c_j, x_j, and d_j but differing values of ε_j) is $CS = E(U_{j*})/\alpha$. Substituting,

$$CS = E(U_{j*})/\alpha = E(W_{j*} + d_{j*})/\alpha = \left(\ln \sum_j e^{-\alpha c_j + \beta x_j} + \sum_j P_j d_j \right)/\alpha \tag{1}$$

Note that $E(W_{j*})$ is the expectation of the maximum value of W_j, which is the familiar log-sum term term in the last line (Small and Rosen, 1981). So the first element in the calculation of CS is the log-sum term calculated on the basis of anticipated utility. The second element is the average difference between actual and anticipated utility.

The average loss from imperfect foreknowledge is

$$\Delta CS = E(U_{j*})/\alpha - E(U_{k*})/\alpha = E(W_{j*} + d_{j*})/\alpha - E(U_{k*})/\alpha = \left(\left[\ln \sum_j e^{-\alpha c_j + \beta x_j} - \ln \sum_j e^{-\alpha c_j + \beta x_j + d_j} \right] + \sum_j P_j d_j \right)/\alpha. \tag{2}$$

This loss is simply the change in log-sum terms from the anticipated attributes to the actual attributes, plus the average difference between actual and anticipated utility. Note that if $d_j \leq 0$ for all j, representing actual utility the same or worse than anticipated, and the inequality is strict for at least one alternative, then the first term (the change in log-sums) is strictly positive and the second term (the loss relative to anticipated utility) is strictly negative. The signs reverse when $d_j \geq 0 \; \forall \; j$ with a strict inequality for at least one alternative. In both cases, the negative term exceeds the positive term in magnitude, since their sum, ΔCS, is strictly negative when $d_j \neq 0$ for at least one alternative. The loss from imperfect foreknowledge is also the gain from obtaining full information about the actual attributes.

These terms can be easily visualized. Fig. 1 depicts a situation in which the actual cost of the good is higher than the anticipated cost.

I use linear demand for one good in order to elucidate the concepts, even though demand for a logit model is nonlinear and for multiple goods simultaneously. The number of people who choose the good is Q^*, where anticipated cost intersects the demand curve. However, consumer surplus is not the triangle AFG as would usually be the case. Rather, consumer surplus is determined by assessing the difference between the demand curve (which gives maximum willingness to pay) and the actual cost. For units of consumption at which the demand curve is above the actual cost, a positive surplus is obtained, as depicted by the blue triangle ABC. For units of consumption at which the demand curve is below the actual cost, a negative surplus is obtained, which is the red triangle CEG: the actual cost exceeds the benefits of consuming these units. Total consumer surplus is the area of the blue triangle minus the area of the red triangle.

This quantity can be readily calculated. The surplus that consumers anticipated receiving is the triangle AFG. For logit models, this area is measured as

$$\left(\ln \sum_j e^{-\alpha c_j + \beta x_j} \right)/\alpha,$$

(*footnote continued*)
to assess welfare when imperfect foreknowledge changes the distribution of terms that are unobserved by the researcher as well as attributes that the researcher observes.

i.e., the log-sum term based on anticipated utility. The difference between actual and anticipated utility for the units purchased is the rectangle BEGF, which for a choice model is $\sum_j P_j d_j$. This rectangle represents a negative since actual utility is worse than anticipated utility in this example (i.e., $d_j < 0$) and constitutes the loss relative to anticipated utility. The area AFG minus area BEGF is equivalent to the blue area ABC minus the red area CEG, which is consumer surplus under the inaccurate anticipation of costs.

The loss due to consumers' imperfect foreknowledge about costs is the red area. The blue area represents people who would have bought the good anyway, even if they had known its actual cost. The red area represents the negative surplus of people who would not have bought the good if they had known its actual cost. The area of this red triangle can be calculated easily as the change in log-sum term under anticipated costs versus actual costs, which is the four-sided area BCGF, minus the rectangle BEGF.

Note that, in this example, all consumers obtain less utility than they thought they were going to obtain. This loss relative to anticipated utility is the area BEGF. However, many of the these consumers would have bought the good anyway, and for these consumers, the actual utility they received for the alternative they chose under their imperfect foreknowledge is the same as they would have received with perfect foreknowledge. Only the consumers who would not have bought the good if they had known its actual cost incur a loss from imperfect foreknowledge. And even for these consumers, their loss from imperfect foreknowledge is less than their loss relative to anticipated utility.

The discussion so far has been in reference to standard logit models, which have fixed coefficients. The concepts are readily generalized for mixed logits, which have random coefficients. Let actual and anticipated utilities be specified as above but with coefficients α and β being random with density $f(\alpha, \beta)$. Consumer surplus under imperfect foreknowledge, and the loss in consumer surplus due to the imperfect foreknowledge, are the same as in Eqs. (1) and (2) above, but integrated over the distribution of the random coefficients:

$$CS = \int \left[\left(\ln \sum_j e^{-\alpha c_j + \beta x_j} + \sum_j P_j d_j \right) / \alpha \right] f(\alpha, \beta)\, d\alpha\, d\beta \tag{3}$$

and

$$\Delta CS = \int \left[\left(\ln \sum_j e^{-\alpha c_j + \beta x_j + d_j} - \ln \sum_j e^{-\alpha c_j + \beta x_j} - \sum_j P_j d_j \right) / \alpha \right] f(\alpha, \beta)\, d\alpha\, d\beta \tag{4}$$

In the following sections, the formulas are applied to two situations: (i) an unanticipated drop in gas prices, and (ii) the privacy policies of video streaming services, which are not generally understood by consumers.

3. Gas price drop

Allcott (2013) used the above formulas to estimate the welfare loss from an "mpg illusion" by which some consumers tend to evaluate fuel efficiency improvements as if savings were linear in mpg, rather than the inverse of mpg. Under this form of miscalculation, the consumer considers, e.g., a 5 mpg improvement in fuel efficiency to reduce operating costs the same amount whether the improvement applies to a base of 20 mpg or 30 mpg, even though the former actually represents over twice as much savings in operating cost as the later. He found that the loss to consumers from this illusion is $3.66 per vehicle purchase, which, given about 16 million purchases per year, comes to an aggregate loss of $59 million per year. He notes that "This amount is large relative to the fixed costs of redesigning fuel efficiency labels, so unless these labels are highly ineffective at debiasing consumers, it seems likely that they could increase welfare."

In the current paper, I borrow Allcott's demand setup to examine another form of imperfect foreknowledge relating to operating costs in vehicle purchase choice. The dramatic drop in gas prices in late 2014 created a distortion in vehicle purchase decisions that is considerably larger than the mpg illusion and has potentially more far-reaching policy implications. The price of gas dropped about 30 percent, and there is no indication that prices will return to their original levels any time soon. This drop has provided an unexpected windfall for the driving public. However, people who bought vehicles prior to the drop made their purchase choices under expectations about future operating costs that ended up being incorrect. In particular, the operating cost difference between gas-guzzlers and gas-sippers was, at the time of the purchase decision, considered to be greater than actually occurred. Some people might have bought larger, more powerful, but less fuel-efficient vehicles if they had known that the additional operating cost was lower than they had anticipated.

This situation provides an interesting application because actual cost is below anticipated cost, which is the opposite of the graph shown above. The same formulas apply, of course. Also, all consumers obtained greater surplus than they had anticipated for whatever choice they made, but there is still a welfare loss from the imperfect foreknowledge of gas prices: with perfect foreknowledge of future gas prices, some consumers could have been even better off buying larger, more powerful vehicles.

To examine this situation, I use Alcott's demand model of new vehicle purchases. Let $c_j(g)$ be the purchase price of vehicle j plus the present value of future operating costs of the vehicle at a gas price of g. Anticipated utility differs from actual utility insofar as anticipated g differs from actual g. Gas for the three years up to mid-2014 averaged about $3.60 per gallon.

Under expectations that these prices would continue, anticipated utility is

$$W_j = -\alpha c_j(3.60) + \beta x_j + \varepsilon_j$$

where x_j captures the observed non-cost attributes of the vehicle. By early 2015, gas price was \$2.50. Actual utility is then

$$U_j = -\alpha c_j(2.50) + \beta x_j + \varepsilon_j$$

such that the difference is $d_j = \alpha(c_j(3.60) - c_j(2.50)) > 0$.

The model considers 213 makes and models of new vehicles. It is calibrated (i.e., alternative specific constants are calculated) so that predicted shares equal the observed market shares of vehicles sales. The price coefficient is set so that the average own-price elasticity is -5.[7] Allcott provides more details on the model.

The gain from the gas price drop, relative to anticipated price, is the quantity $\sum_j P_j d_j$, where P_j is the probability of choosing vehicle j. Based on the calibrated model, this quantity is \$5,125 per vehicle purchase, which constitutes \$82 billion per purchase-year in aggregate. Lower gas prices do indeed provide a large benefit to consumers. However, consumers would have been even better off if they had known, when choosing a vehicle, that gas prices were going to drop. This loss comes to \$72 per vehicle purchase, or \$1.15 billion per purchase-year.

This loss is considerably greater than that due to the mpg illusion. And the policy implications are disconcerting. If consumers had bought with perfect foreknowledge of future gas prices, the average mpg of new vehicles would have been lower by 0.60 than actually occurred. By standard concepts of market intervention, this distortion, when viewed alone, could be considered a justification for lowering the Corporate Average Fuel Efficiency (CAFE) standards. In my opinion, externalities (including climate change) justify maintaining and raising the current CAFE standards. However, consumers' imperfect foreknowledge about future gas prices when purchasing a vehicle needs to be considered in the analysis.

4. Streaming video services

Some video streaming services share their customers' data with third parties, e.g., for targeted advertising of other products and services. Butler and Garrett (2015) investigate consumers' preferences in relation to data sharing by video streaming services. They report that most consumers do not know that their provider, or any providers, collect customers' usage or demographic information, much less that the providers might share the data with third parties. The question arises: what is the loss to consumers, when choosing among providers, of thinking that the providers do not share their data when in fact some or all of them do?

Butler and Glasgow conducted a standard conjoint analysis of factors that affect consumers' choice among video streaming services. Eligible participants in the study were people who currently purchased a streaming service or were considering doing so within the next year. A mixed logit model was estimated on the choices in conjoint exercises, with the coefficient of price specified as log-normally distributed and the coefficients of the other attributes specified to be normally distributed.

Table 1 gives their estimated mixed logit model. The data sharing practice of each service was specified as taking three possible levels: (1) no sharing, (2) sharing of customers' usage data in aggregate form with no personal information, and

Table 1
Choice of video streaming service.

From Butler and Garrett (2015)

	Mean	Std.Dev.	Average WTP
Share usage	−0.056	0.114	−0.85
Share usage and personal	−0.366*	0.708*	−5.52
Price (log of coefficient)	−2.095*	1.118*	
More movies	0.435*	0.704*	6.62
More TV	0.009	0.735*	0.10
Commercials	−0.418*	0.869*	−6.41
Fast Content	0.534*	0.738*	8.08
No service	−2.944		

Significance not calculated for average WTP.
* Significance at 95 percent confidence level.

[7] For a logit model, the own-price elasticity of demand for vehicle j is $\epsilon_j = -\alpha(1 - P_j)r_j$, where P_j is the probability of choosing vehicle j and r_j is its price. The value of α that gives an average elasticity of -5 is then 5 divided by the mean of $(1 - P_j)r_j$. The procedure is generalized for nested logits.

Table 2
Sample shares.

Netflix	0.29
Amazon Prime	0.05
Hulu	0.04
Netflix+Amazon Prime	0.12
Netflix+Hulu	0.07
Amazon Prime+Hulu	0.01
Netflix+Amazon Prime+Hulu	0.07
None	0.36

(3) sharing of customers' usage data and personal information. The model includes variables for the two forms of sharing relative to not sharing. As shown in Table 1, sharing of usage data in aggregate form does not enter significantly: neither the mean nor standard deviation of the coefficient is statistically different from zero at usual levels of confidence. The point estimates imply that the average willingness to pay (WTP) to avoid having their usage data shared in aggregate form is only 85 cents per month.

Consumers are considerably more concerned about the sharing of their usage data coupled with personal information. The mean and standard deviation are statistically significant, and the point estimates imply an average WTP of $5.52 per month to avoid having their usage and personal data shared.[8]

The other attributes in the model also have interesting implications. People seem to greatly dislike the prospect of having commercials shown before the video: average WTP to avoid commercials is $6.41. “Fast content” means that TV shows and movies are available on the service shortly after they are aired or released. This attribute is estimated to be worth, on average, $8.08. The conjoint exercises included the option of not purchasing any service, and the model estimates a “no service” constant for this option.[9]

The survey asked which services, if any, the respondent had purchased, asking explicitly about Netflix, Amazon Prime, and Hulu. Table 2 gives the share of survey consumers who have each service or combination of services. Many consumers choose a combination of services. And, as becomes relevant in our discussion below, Hulu is chosen most often in combination with another service: only 21 percent of Hulu subscribes have only Hulu.

As stated above, most consumers did not know that services shared data. Our goal is to estimate the loss to consumers of this lack of knowledge, which also constitutes the gain that would be obtained if consumers were informed of the sharing practices.[10]

In both scenarios, consumers are assumed to believe that none of the services share their data. The scenarios differ in what is assumed to actually occur. Under scenario 1, Hulu is assumed to share its customers' usage and personal information with third parties, but the other services are not. The reason for choosing Hulu is pedagogic: it illustrates an important concept, as shown below. Under scenario 2, all services are assumed to share their customers' usage and personal information.

The alternatives available to consumers are specified to be those in Table 2: Netflix, Amazon Prime, Hulu, each combination of them, and no service. The price of Netflix and Hulu is set at $7.99 per month, and the price of Amazon Prime is $6.85 per month. Anticipated utility is calculated with none of the services sharing data. The model is calibrated so that forecasted shares equal the sample shares in Table 2.

4.1. Scenario 1: Hulu shares usage and personal data

We first consider a situation in which Hulu shares usage and personal data but the other providers do not. Hulu had about 6 million subscribers at the time of the survey. The model in Table 1 indicates that consumers, on average, are willing to pay $5.52 per month to avoid having their usage and personal data shared. A naive calculation would suggest that the aggregate loss from Hulu sharing its customers' data is $5.52 times 6 million=$33 million per month. However, this calculation misses several important concepts. First, in a mixed logit model, the average WTP among people who chose a given option (in this case, Hulu) need not be the same as the average WTP among all consumers. Stated equivalently, the

[8] The estimated standard deviation indicates that some consumers prefer to have their usage and personal data shared, perhaps because they value the information that they might obtain from third-party advertisers.

[9] Butler and Glasgow's model also includes an interesting feature of estimating the share of survey respondents whose WTP to avoid sharing data cannot be calculated from their conjoint responses. As the authors explain, these respondents might have protested the concept of data sharing in their answers to the conjoint exercises, or might have WTP to avoid sharing that is larger than the price points used in the exercises. For the current analysis, I omit these respondents.

[10] Providers usually disclose their sharing practices in policy statements to consumers. Of course, a disclosure statement does not necessarily make the consumer informed, since few people read or understand these detailed statements, as evidenced by the small share of consumers that Butler and Glasgow found knew that their data were collected. The goal of the exercise is not to place blame for the losses, but rather to determine the magnitude of the losses. Also, there is a question of whether sharing data actually hurts consumers if the consumers do not know it is happening. Customers can be hurt by receiving more popups, emails, and other forms of targeted advertisement from third parties, even if the customer does not know why the advertisements arrive. For the sake of our illustration, I assume that consumers, on average, do indeed lose utility from the data sharing.

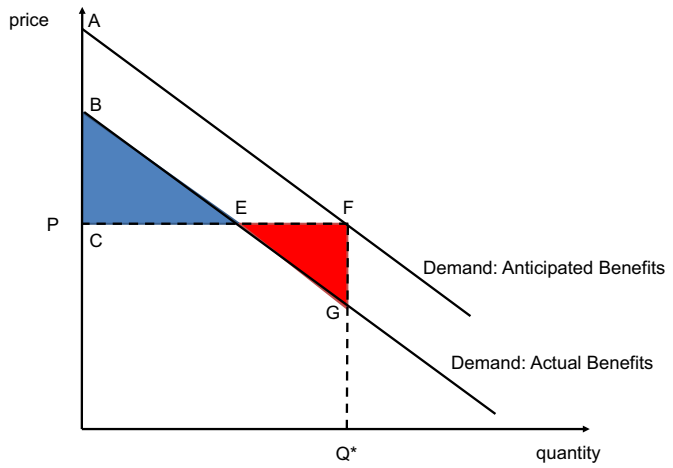

Fig. 2. Consumer surplus under imperfect foreknowledge about sharing of data.

conditional distribution of WTP is not the same as the unconditional distribution. Second, some Hulu subscribers might have chosen Hulu even if they had known that Hulu shared their data. These subscribers incurred no loss from their lack of information. Third, consider the Hulu subscribers who would not have chosen Hulu if they had known their data would be shared. The loss to these consumers due to their misunderstanding about the sharing practices is not their entire WTP to avoid sharing, since they obtain some benefits from the service. Application of the above formulas accounts for all three of these issues.

The loss that consumers experienced relative to their anticipated utility is $\sum_j P_j d_j$ integrated over the distribution of random coefficients. This quantity comes to \$54 million per month in aggregate, which is \$9 per Hulu subscriber per month. This amount is considerably greater than the naive calculation of \$5.52 per Hulu subscriber. The difference is due to the fact that the average WTP among Hulu subscribers is considerably higher than the WTP for all consumers. As shown in Table 2, Hulu is mainly purchased in combination with other services. The combination services cost more than the single services or no service. And consumers with a small coefficient for cost (i.e., low in magnitude) are more likely to choose the combination services. Since WTP is defined as the attribute coefficient divided by the cost coefficient, these consumers have high WTP.

\$54 million is the amount by which the surplus that consumers actually obtained falls below the surplus that they thought they would obtain. However, some of these consumers would have bought Hulu even if they had known that their data would be shared. Their misunderstanding of its sharing practice did not affect their choice, and hence did not affect the surplus they actually obtained. The loss due to consumers' not knowing the sharing practice is calculated by Eq. (2) to be \$24 million per month, which is \$4 per Hulu subscriber per month. Only 16 percent of Hulu subscribers would not have chosen Hulu if they had known their data would be shared. These customers generally have higher WTP to avoid sharing than other customers, which is the reason that \$24 million is more than 16 percent of \$54 million.

Fig. 2 depicts the situation for one value of the random coefficients. It differs from Fig. 1 because the imperfect foreknowledge is about a non-cost attribute of the product, rather than cost. The higher demand curve represents the anticipated benefits of buying Hulu and the lower one represents the actual benefits. The number of people who subscribe to Hulu is Q^* under their incorrect belief about sharing. Consumer surplus, given the actual utility from Hulu, is the blue triangle BCE (a positive surplus) minus the red triangle EFG (negative surplus). This combined surplus is less than consumers had anticipated receiving, by the area AFGB. Consumers who still obtain a positive surplus would have purchased Hulu even with full information. The loss due to imperfect foreknowledge about sharing is the red triangle EFG. Our calculation of \$24 million in loss is the red area integrated over the distribution of coefficients.

4.2. Scenario 2: all providers share usage and personal data

Under this scenario, all providers share their customers' data, but, as in scenario 1, customers do not know this when choosing. I assume 20 million subscribers, which is consistent with 6 million Hulu subscribers and the sample shares in Table 2. The naive calculation of loss is average WTP of \$5.52 per month times 20 million subscribers, which is \$110 million per month. As before, this calculation misses the fact that people who choose to subscribe might have a different WTP on average than the average WTP in the population, and that subscribers might have chosen to subscribe even if they knew their data were being shared. The reduction in surplus relative to what consumers thought they would obtain is \$145 million, which is higher than the naive calculation because subscribers have a higher WTP on average than non-subscribers. The loss due to the imperfect foreknowledge about sharing is \$18 million per month. This is a very small share of the loss

relative to anticipated utility ($145 m), because over 90 percent of subscribers would have still subscribed if they had known the sharing policies. Since all services share their customers' data, the only way a customer can avoid having their data shared is to not subscribe to any service. Few subscribers would have chosen to do without any service if they had known the sharing policies.

Importantly, the loss due to imperfect foreknowledge is considerably smaller under scenario 2, when all services share their customers' data, than under scenario 1, when only one service does. When only one service shares its data, consumers could have avoided the sharing by choosing the other services. But when all services share their data, the only way for a customer to avoid the sharing is to do without any video service. This difference in substitution options explains the seeming paradox that loss from imperfect foreknowledge is lower when the same mistaken anticipation applies to more alternatives. At an extreme, if a consumer's anticipations are equally incorrect for all alternatives (i.e., $d_j = d_k \ \forall\, j, k$), then imperfect foreknowledge has no effect on the consumer's choice and the consumer is not harmed by it. This circumstance is probably rare, though. In the more usual situation, with anticipations about attributes affecting different alternatives differently, the consumer can lose utility from choosing the alternative that maximizes anticipated utility rather than actual utility.

References

Allcott, Hunt, 2013. The welfare effects of misperceived product costs: data and calibrations from the automobile market. Am. Econ. J.: Econ. Policy 5 (3), 30–66.

Butler, Sarah, Garrett, Glasgow, 2015. The Value of Personal Information to Consumers of Online Services: Evidence from a Discrete Choice Experiment. Working Paper, NERA, Inc., University of California, Santa Barbara.

Delavande, Ademine, 2008. Pill, patch, or shot? Subjective expectations and birth control choice. Int. Econ. Rev. 49 (3), 999–1042.

Jensen, Robert, 2010. The perceived returns to education and the demand for schooling. Q. J. Econ. 125 (2), 515–548.

Kahneman, Daniel, 1994. New challenges to the rationality assumption. J. Inst. Theor. Econ. 150 (1), 18–36.

Manski, Charles, 2004. Measuring expectations. Econometrica 72 (5), 1329–1376.

Schmeiser, Steven, 2014. Consumer inference and the regulation of consumer information. Int. J. Ind. Organ. 37, 192–200.

Small, K., Rosen, H., 1981. Applied welfare economics of discrete choice models. Econometrica 49 (1), 105–130.

van der Klaauw, Wilbert, Kenneth, Wolpin, 2008. Social security and the retirement and savings behavior of low-income households. J. Econom. 145 (1–2), 21–42.