Economic Value of Location-based Big Data: Estimating the Size of Japan's B2B Market (*)

Teppei KOGUCHI (**)
Shizuoka University, Japan

Toshiya JITSUZUMI
Kyushu University, Japan

Abstract: The rapid development of data handling and statistical analysis has increased international interest in using "big data" for multiple applications. However, many stakeholders are hesitant to take the first step or to make an initial investment because the projected market size is still unknown, despite the high probability of a positive feedback loop once it takes off. This paper provides a basic foundation for further discussion by empirically estimating the size of the B2B market for location-based data in Japan. The results of a questionnaire survey following a conjoint method suggest that location-based data currently produced and stored by Japanese mobile operators can generate as much as 195 billion yen a year, which is equivalent to 11% of the annual operating profits of the Japanese mobile sector.

Key words: big data, location-based data, dataset market.

In recent years, the rapid development of data handling and statistical analysis has increased international interest in using "big data" for multiple applications, a trend facilitated by recent advancements in information and communications technologies (ICTs) that enable the easy collection and dissemination of vast amounts of data. Accordingly, many governments consider these data as the "new oil" and are seeking ways to efficiently use them to enhance social welfare, while businesses are striving to use them to enhance operational efficiency and to monetize them effectively.

(*) An earlier version of this paper was presented at the 25th European Regional International Telecommunications Society Conference, Brussels, Belgium, June 22-25, 2014.

(**) We would like to thank our two anonymous reviewers for their valuable comments. This research was supported by JSPS KAKENHI Grant Number 25380304. We are also grateful for a grant from KDDI Research Institute and the support of the JIPDEC.

As was illustrated by the discussions relating to the IT productivity paradox [1], using this new oil efficiently and turning big data into an indispensable input for firms require the coordinated efforts of various stakeholders. Consumers must understand the benefits and risks of the widespread utilization of personal data; data-generating firms must develop cost-effective methods for data handling, accompanied by robust privacy safeguards; data-using firms must modify their business models in order to realize big data's economic potential; and policymakers must introduce a new privacy framework that accommodates the rapid pace of ICT development and the welfare-enhancing yet privacy-threatening potential of the big data ecosystem. The eventual widespread use of big data is expected to increase productivity in many industries and grow the overall economy.

This big data ecosystem involves indirect network externalities between data-generating and data-using firms. The larger the data-using group becomes, the more big data will be provided by data-generating firms for profit, which will then attract more user firms. Given the interrelatedness among the players, however, this positive, or welfare-enhancing, feedback will not ignite unless the number of players exceeds the critical mass. Therefore, it may be economically justified to provide government support in order to surpass the critical mass level of utilization.

However, at least three obstacles impede this coordination. The first is the lack of proper understanding of this new economic resource. For example, we still lack a proper definition of "big data," causing many to see it as an ephemeral "buzzword" unworthy of serious consideration. Another obstacle, one widely observed in Japan, is that a B2B market for big data has not yet been developed. Except for the rare occasions when a company is fortunate enough to know somebody who can provide them with necessary big data, user firms are generally limited to those who can generate big data by themselves. Thus, managers who want to use big data for, say, marketing must design their own data-collection method from scratch; those who cannot are unable to realize big data's benefits. Last but not least is the legal obstacle. In Japan, under the Personal Information Protection Act, companies do not have full discretion in their use of big data. User firms must thus be cautious when using data containing information covered by the act, increasing the marginal cost of big data use and deterring its penetration. The ambiguity of the definition of "personal

[1] Related discussions are summarized in JITSUZUMI (2005).

information" aggravates the problem, as became clear in the case of the East Japan Railway Company (JR East). In June 2013, JR East announced that it would begin to provide movement data on its train passengers, collected by its e-tickets and automatic ticket checkers, to a third-party firm. JR East claimed that the dataset did not include passengers' "personal information" (such as names) and was therefore not expected to cause problems under the act. In reality, however, many experts doubted the legality of the plan, and JR East thus had to drop it.

From an economic perspective, overcoming these obstacles and realizing the potential value of big data require that market players understand big data's value and that investment be lured into this market sector. The authors believe that the B2B market can contribute significantly to the realization and dissemination of big data's value. The fact that each dataset has vast value-generation possibilities that spread much farther than data-generating firms can integrate into or exploit through their current business operations makes it difficult for them to enjoy the resource's full potential. Without the opportunity to collect value by selling data to other user firms (thus internalizing the spillover benefits), investments in big data generation are likely to be suboptimal, and the size of the ecosystem will become inefficiently small.

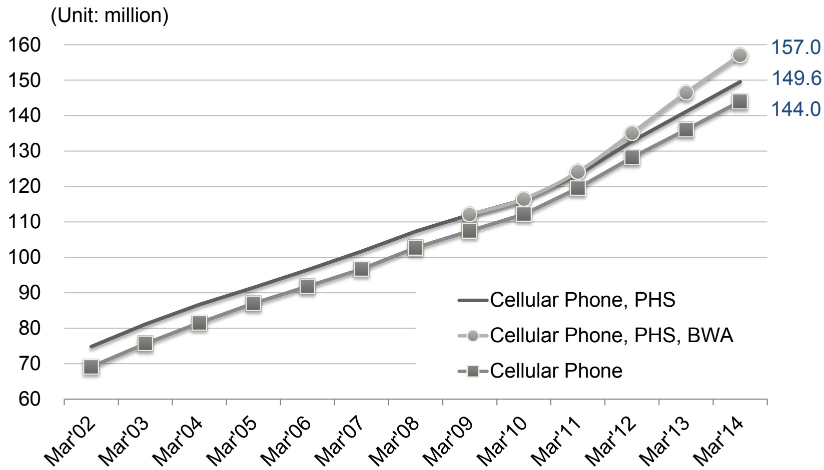
Several attempts have been made to estimate the economic significance of big data. For example, their impacts on overall productivity, innovation value, and business creation are analyzed in BCG (2012), CEBR (2012), General Electric (2013), and MANYIKA *et al*. (2011), while their firm-level impacts are analyzed in LAVALLE *et al*. (2011) and RUSSOM (2011). In addition, CHEN *et al*. (2012) survey the extant studies from a management perspective. In a study relevant to the size of the B2B market for location-based big data, the Japan Institute for the Promotion of Digital Economy and Community (JIPDEC) found that the unit price for such data was between one and five yen (2012). However, this figure cannot be considered fully accurate and useful for two reasons. First, JIPDEC's result was based on interviews with 15 experts and must thus be tested statistically before being used for policy making. Second, JIPDEC failed to investigate whether the dataset's contextual details (e.g., whether it included the gender or age of the individual data points) could have influenced its economic value.

This paper seeks to overcome these flaws by empirically estimating the size of the B2B market for big data in order to provide the basic information necessary for a discussion of the coordination among stakeholders. The rest of this paper is organized as follows. The next section explains the empirical estimation of the monetary impact of variations in location-based big data. Using these estimates, we calculate the size of the B2B market in the third section. Finally, the fourth section concludes the paper.

■ Value of location-based data

Figure 1 shows that the number of subscriptions is steadily increasing in the Japanese mobile market: there were around 70 million in March 2002, over 100 million in March 2007, and the number continues to increase. The Ministry of Internal Affairs and Communications (MIC) claims that the penetration rate was 101.7% in March 2014, meaning that, on average, each Japanese individual has more than one mobile connection.

Figure 1 - Trend in the number of mobile phone subscriptions



Note: PHS is a Personal Handy-phone System, and BWA is Broadband Wireless Access

Source: Created on the basis of MIC (2014)

In recent years, the market has experienced the rapid popularization of multifunctional mobile phones, or "smartphones." Unlike "non-smart" handsets, smartphones can be equipped with many sensors, including cameras, acceleration meters, compasses, microphones, and GPS [2]. Thus,

[2] In 2007, the MIC obliged all mobile phone terminals to be equipped with GPS to strengthen the emergency report function.

as smartphone penetration advances and as mobile usage steadily increases, mobile phone operators have a huge opportunity to collect massive amounts of data that can provide valuable information about subscribers' behaviors. Moreover, the fierce competition among mobile operators and their massive investments have improved mobile services' population coverage, which reached 99.97% in November 2013. This has improved mobile operators' data-collection capability even more, allowing them to obtain data on the whereabouts of their subscribers seamlessly and without time lags. Thus, in this and subsequent sections, we focus on the mobile phone business' most promising data – the location-based big data generated by the Japanese mobile operators themselves. These data can benefit several industries in many ways; however, for simplicity's sake, the authors focus on their usage by "mobile content firms," which use mobile handsets to provide and/or distribute services, and estimate the extent of the willingness to pay (WTP) in order to measure the potential B2B market for them.

The estimation is based on a questionnaire survey conducted from November 2013 to January 2014 on a population of 1,048 potential user firms [3]. Of these, 1,000 were selected through the following method: first, we selected 2,373 candidates using a keyword search [4] of each firm's business contents listed in the database of Teikoku Databank, Ltd.; then, we selected 1,000 from among those with the highest annual revenue. Additionally, 48 were selected from the 50 firms on the membership lists of The Consortium for the Promotion of Next-Generation Personal Service [5] and the g-Contents Exchange Promotion Association [6]. These groups are comprised of firms that are interested in big-data utilization. Due to this selection procedure, our estimation results may be biased if directly applied to the whole economy. However, since the big-data market is in its infancy and a high level of literacy cannot be generally expected from Japan's industries, we believe it beneficial to focus on the most promising market. The questionnaires were

[3] Our questionnaire survey was conducted during a research project led by the Ministry of Economy, Trade and Industry (MITI). Thus, this paper serves as a follow-up of the MITI's 2013 presentation at the OECD's Working Party on the Information Economy, which studies developments, trends, and policy implications surrounding digital content.

[4] The keywords used were "social networking," "social game," "site," "content," "game," "Android," "App Store," "application mobile," "mobile phone," "mobile handsets," "smartphone," and "iPhone."

[5] https://www.coneps.jp/

[6] http://www.g-contents.jp/

distributed by mail along with explanatory material, and responses from 176 firms (16.8% [7] of the population) were collected for further evaluation.

Figure 2 - Difference in coverage

The values of big data have not sufficiently materialized and are expected to vary depending on the data contents, data size, granularity, and area coverage; thus, the authors followed a conjoint method [8] to construct the questionnaire and empirically estimate the WTP figures for individual big data attributes. First, the questionnaire asked respondents to assume the following situations:

---

[7] Since the questionnaires were distributed to firms rather than to consumers or households, this collection rate is not particularly low.

[8] The conjoint method uses stated preference data for the respondents in a hypothesized situation and is thus particularly useful for estimating the market value of goods or services that are too new to have real market transaction data. For more detail, please refer to LOUVIERE *et al*. (2000, Chapters 4-7).

- Mobile phone operator "A" announces that it is now selling datasets containing location-based data on its subscribers to business users.
- The data include pseudo ID, measurements of time/day, coordinates of longitude and latitude, elevation, and degrees of granularity and is collected every five minutes from individual handsets automatically and stored for a year.
- The dataset can be provided accompanied by demographic information (such as sex and age group) on the individual data points.
- All data are pseudonymized; thus, users cannot track back to individual customers but can track individual movements for one year.
- The datasets contain a year's worth of data on one million subscribers randomly selected from A's nationwide user base (for the nationwide case) or one year's worth of data on 100 thousand subscribers randomly selected from one of A's regional markets in Hokkaido, Tohoku, Kanto, Chubu, Kinki, Chugoku, Kyushu, or Okinawa (for the regional case, see Figure 2).

Questionnaire respondents were asked to choose their preferred option from among four alternatives, including "Do not purchase." The remaining three alternatives were different from each other concerning at least one of the four attributes (see Table 1). The lineups and the levels of the individual attributes were determined based on the findings of the JIPDEC (2014) and preliminary discussion with six big data experts who had participated in the METI working group (see Table 2). Here, "sex" means with or without information about the sex of handset holders; "age" means with or without age information (by decade) of handset holders; "mesh" represents the granularity at which individual location-based data are aggregated; and "price" is a one-time payment for a dataset. The number of alternatives is reduced to 16 through orthogonal design to lower the stress of answering the questionnaire as far as possible while maintaining its statistical validity. Each sheet contains 10 questions, the first five for the nationwide case and the remaining five for the regional case [9]. The questionnaire was split into nationwide and regional cases for three reasons. First, the field of expected application for the nationwide dataset may not be the same as that for the regional dataset; thus, their values may differ considerably. In addition, the expected users of the nationwide dataset may differ from those of the

[9] To verify the respondents' consistency, the fifth and tenth questions have the same alternatives as the first and the sixth but in a different order. Therefore, 12 alternatives are presented for individual respondents. To cover all 16 alternatives, respondents were divided into four segments and were asked to answer different alternative combinations.

regional dataset, which may also affect the comparability of the values. Finally, since no linear or other constant relationship between the dataset's economic value and its data density and/or the size of its coverage has been proven, the value of the nationwide dataset relative to that of the regional one cannot be predicted.

Table 1 - An example of four alternatives

		Alternative 1	Alternative 2	Alternative 3	Alternative 4
Feature of the base dataset	Mesh size	$5km^2$	$3km^2$	$500m^2$	Do not purchase
	Price	1 million yen	2 million yen	4 million yen	
Addition of demographic data [optional]	Sex	Without	Without	With	
	Age	Without	With	With	

Table 2 - Conjoint attributes and their levels

Attributes	Levels			
Sex	With	Without		
Age	With	Without		
Mesh size	$100m^2$	$500m^2$	$3km^2$	$5km^2$
Price				
nationwide case	1,000,000 yen	2,000,000 yen	4,000,000 yen	8,000,000 yen
regional case	200,000 yen	400,000 yen	1,000,000 yen	2,000,000 yen

The descriptive statistics for our data are summarized in Table 3. It shows that most of the respondents who produced valid responses have headquarters in Tokyo, indicating that viable demand may be concentrated in Tokyo, where many firms market their products on a nationwide basis. If this is so, a regional location-based dataset will be difficult to sell.

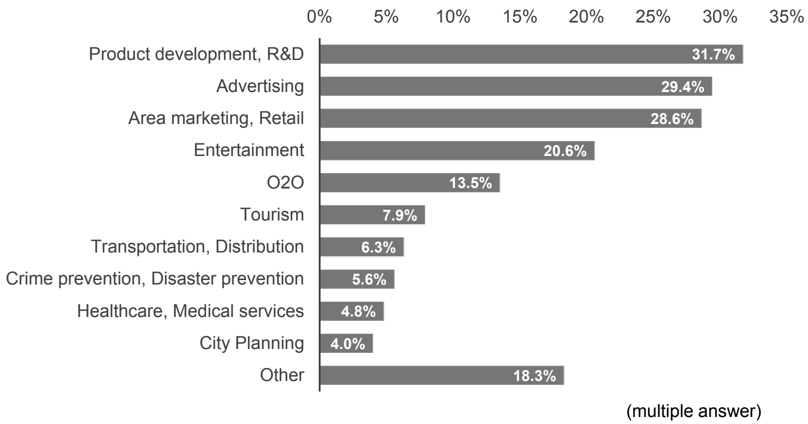
To determine whether the responding firms had the data-analysis skills required to use the location-based big data, we asked them to name the areas where they were currently using their skills. The result, shown in Figure 3, indicates that "product development and R&D," "advertising," and "area marketing and retail sales" are the areas for which most data analytics are used. These areas are not where location-based data can achieve their potential (except for area marketing), showing that many respondents may lack proper skills necessary for utilizing location-based data. Thus, they were not interested when offered a set of location-based data. This could explain why there were so many "all negative responses" in our data.

Table 3 - Status of the responding firms

	nationwide	regional
N	176	
blank responses	11	12
inconsistent responses	41	36
consistent responses		
all negative responses	104	93
mixed responses	21	36
(of mixed responses) place of headquarters		
Tokyo	12	22
Other	9	14

Note: The "blank responses" category comprises individuals who did not provide answers. The "inconsistent responses" category comprises individuals whose capricious or inconsistent answers become apparent when the first and sixth questions are compared to the fifth and tenth, which offer the same alternatives but in a different order in each questionnaire. The "all negative responses" category comprises individuals who chose "no purchase" for all questions. The "mixed responses" category comprises individuals who show both negative and positive purchase intention.

Figure 3 - Areas of business where data analytics are currently used



Source: Created on the basis of JIPDEC (2014, p. 33)

The estimation is conducted following random utility theory, where the utility of user *i* from choosing alternative *j* is described as a combination of representative utility V_{ij}, and an error ε_{ij} (Equation 1), and the probability of choosing alternative *k* is described as in Equation 2. Here, *f* is the joint distribution function of the error term, and *I* is an index function that equals one if the inside of the following parenthesis is true and zero otherwise:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad [1]$$

$$\begin{aligned} P_{ij} &= \mathrm{prob}\left(U_{ik} > U_{ij}, \forall j \neq k\right) \\ &= \int I\left(\varepsilon_{ij} - \varepsilon_{ik} < V_{ik} - V_{ij}, \forall j \neq k\right) f\left(\varepsilon_i\right) d\varepsilon_i \end{aligned} \quad [2]$$

Since the independence from irrelevant alternatives (IIA) assumption is not statistically supported [10], a random parameter logit model (RPLM) is employed, where the normal distribution is assumed for every parameter (β) except one for a data granularity and one for a price variable. The functional specifications employed are shown as Equations 3 and 4.

$$P_{ij} = \int \frac{\exp(\boldsymbol{\beta}_i'\mathbf{X})}{\sum_j \exp(\boldsymbol{\beta}_i'\mathbf{X})} g(\boldsymbol{\beta}|\boldsymbol{\theta}) d\boldsymbol{\beta} \quad [3]$$

$$\begin{aligned} U_{ij} = &\left(\beta_{sex} + \sigma_{sex} v_{sex}\right) D_{sex} + \left(\beta_{age} + \sigma_{age} v_{age}\right) D_{age} \\ &+ \beta_{invmesh} InvMech + \beta_{price} Price + \varepsilon_{ij} \end{aligned} \quad [4]$$

Where *InvMesh* is an inverse of measurement mesh (in km^2) and equals 0 when respondents do not purchase any datasets; *v* is the individual specific heterogeneity with a mean of zero and a standard deviation of one; *εij* is an error term; and *βs* and σs are parameters.

To focus on the most promising user firms (as mentioned), the estimation uses data only from the firms with mixed responses. This results in 84 samples from 21 firms in the nationwide case and 141 samples from 36 firms in the regional case. Tables 4 and 5 show the estimated results. The estimated parameters (*βs*) for the nationwide case have the theoretically expected signs and are statistically significant, except for β_{age}. In the regional case, β_{sex}, though statistically significant, has an unexpected negative value, and β_{age} as well as β_{price} turn out to be insignificant. Divided by β_{price}, these parameters can be converted into a monetary value, or WTP figures, and can be used to derive the maximum price that an average respondent may be willing to pay for the proposed dataset (see Table 6) [11].

[10] Details of the IIA test results are available upon request.

[11] Since the price parameter in the regional case is not statistically significant, WTPs for the regional dataset are given for reference only and must be interpreted cautiously.

Table 4 - Estimated parameters for the nationwide case

N		84		
Log likelihood		-72.685		
Restricted log likelihood		-116.449		
McFadden Pseudo R^2		0.376		
Variable		**Coefficient**	**Standard Error**	**p-value**
D_{sex}	$\hat{\beta}_{sex}$	0.981	0.588	0.095
	$\hat{\sigma}_{sex}$	0.659	2.225	0.767
D_{age}	$\hat{\beta}_{age}$	0.850	0.715	0.234
	$\hat{\sigma}_{age}$	0.118	0.862	0.891
$InvMesh$	$\hat{\beta}_{invmesh}$	0.011	0.005	0.041
$Price$	$\hat{\beta}_{price}$	-0.009	0.002	0.000

Table 5 - Estimated parameters for the regional case

N		141		
Log likelihood		-126.253		
Restricted log likelihood		-195.468		
McFadden Pseudo R^2		0.354		
Variable		**Coefficient**	**Standard Error**	**p-value**
D_{sex}	$\hat{\beta}_{sex}$	-7.817	8.681	0.368
	$\hat{\sigma}_{sex}$	4.855	6.097	0.426
D_{age}	$\hat{\beta}_{age}$	0.913	0.727	0.210
	$\hat{\sigma}_{age}$	1.609	1.474	0.275
$InvMesh$	$\hat{\beta}_{invmesh}$	0.013	0.003	0.001
$Price$	$\hat{\beta}_{price}$	-0.002	0.004	0.578

For the nationwide case, although the direction of influence from each specification matches our ex-ante expectation, the monetary impact of decreasing data granularity seems excessive [12]. However, the highest price a mobile operator can ask for a nationwide dataset with the finest granularity is within the range of what JIPDEC (2012) found for the price of location-

[12] This is probably due to the assumed linear relationship between granularity and utility level. The authors tried other functional specifications that allow non-linear relationships, but the main parameters were not significant.

based data [13]. In addition, these figures show that adding demographic information on each data point would double the value of the whole dataset.

Table 6 - Derived WTPs for location-based datasets

The nationwide case				
	No demographic data	with sex	with age	with sex and age
100m^2 mesh	¥1,199,060	¥2,299,749	¥2,152,833	¥3,253,522
500m^2 mesh	¥47,962	¥1,148,652	¥1,001,735	¥2,102,425
3km^2 mesh	¥1,332	¥1,102,022	¥955,105	¥2,055,795
5km^2 mesh	¥480	¥1,101,169	¥954,253	¥2,054,942

The regional case (reference)		
	No demographic data	with age
100m^2 mesh	¥5,689,750	¥9,606,925
500m^2 mesh	¥227,590	¥4,144,765
3km^2 mesh	¥6,322	¥3,923,497
5km^2 mesh	¥2,276	¥3,919,451

■ Size of the B2B market

The assessment in the previous section allows us to estimate the size of the B2B market for location-based big data that are gathered and stored by the Japanese mobile operators. Hereafter, we shall focus on the nationwide dataset with a data collection granularity of 100m^2 mesh.

First, based on the share of the valid responses of our questionnaire analysis, we estimate that the number of potential buyers of the nationwide location-based dataset in the mobile content sector is 16.8% of 2,373 firms, or 398.7 firms [14]. If this estimation is correct, the overall size of the market may range from 72 billion yen to over 195 billion yen, depending on what kind of demographic data are added (see Figure 4).

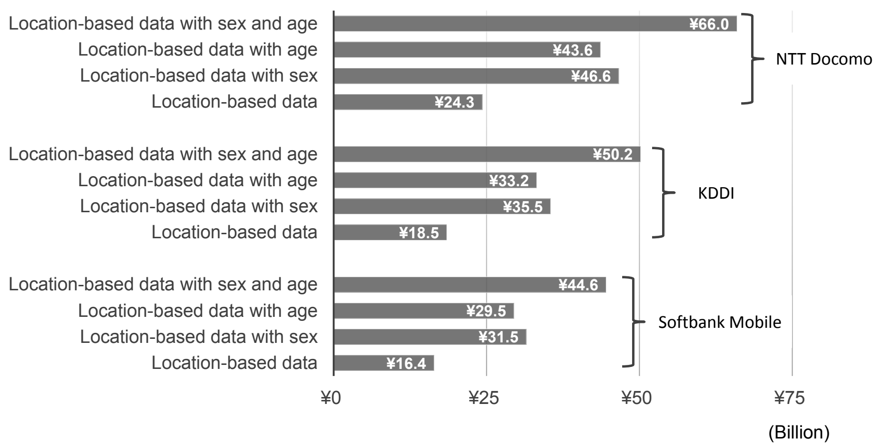
[13] The JIPDEC (2012) estimate for a unit of data is between one and five yen, which makes a dataset of one million data points cost one to five million yen.

[14] Twenty-one divided by 125, or 16.8% is the ratio of mixed responses to all consistent respondents. Moreover, according to Teikoku Databank, 2,373 firms belong to the mobile content sector (for the detailed respondent configuration, see Table 3).

Figure 4 - Potential market size of the estimated dataset market

Further, if we can assume a linear relationship between the value of big data and the number of data points, the value of big data owned by the Japanese mobile operators with 150 million subscribers is about 150 times the estimates shown in Figure 4. If a volume discount is offered in the big-data B2B market, the monetary value increases more slowly than do the data points involved. On the other hand, if a combination of individual data can generate some positive value, the monetary size of the dataset increases faster than does the number of data included. Therefore, we believe that the "linear relationship" assumption adopted here is modest and reasonable.

Figure 5 - The estimated values of mobile operators' nationwide big data



Given the share of the mobile market, the potential values of the location-based datasets owned by each mobile operator can be calculated as shown in Figure 5. As mobile operators collect these location-based data during their daily operations, using these valuable data assets requires no significant additional investment for data generation. Based on these estimates, we can infer that Japanese mobile operators have a latent property that can generate as much as 2.7% to 10.8% of their operating profits annually [15] [16].

■ Conclusion

This study finds that the location-based data collected and stored in the Japanese mobile sector have a potential value as high as 195 billion yen annually, or about 11% of the sector's operating profits. It is important to mention that the overall impact of the B2B market's size is even larger because our study focuses only on the market's most promising segments. We hope our estimates will encourage firms to start thinking about using big data and motivate them to invest more in this area; this will help establish a competitive B2B market, enhancing the efficiency of the overall economy. Further, an accumulation of analyses, such as this study, will help cultivate an understanding of big data's commercial value and introduce an economic perspective in discussions regarding updating privacy policies. It is also believed that the estimated value of bundling demographic information with location-based data will encourage mobile operators to construct an optimal compensation scheme for end users and rationalize opt-in/opt-out conditions.

From the consumers' perspective, this study confirms the huge profit potential of the demographic information they have shared with mobile operators without any monetary compensation. If the results of this analysis, and of similar studies, were widely known, ordinary users would be more

15 The operating profits were 819.2 billion yen for NTT DoCoMo (FY2013), 419.2 billion yen for the mobile section of KDDI (FY2012), and 609.0 billion yen for Softbank Mobile (FY2013).

16 Although data generation does not require additional investment, some sales expenditures are inevitable. In addition, mobile operators may have to incur user education costs in order to build the location-based data market from scratch. However, because their business reputation in the Japanese market can minimize the need for sales promotion and as the learning capacity of potential user-firms may eliminate the need for educational efforts, it is safe to assume that such costs would not gravely impact our current estimates.

informed about the costs/benefits of accepting mobile operators' privacy policies and would rationalize their own behaviors, thus increasing the overall efficiency of the emerging big data ecosystem.

The commercial estimates from this study can also help governments design an appropriate policy package to foster healthy development of the big data ecosystem.

It needs to be emphasized that every aspect of this analysis is highly dependent on the current technological situation, general economic prospects, and the consumers' overall ICT literacy. Therefore, similar studies should be conducted periodically in order to verify whether the implications derived from our study remain valid.

Finally, future research faces several challenges. First, an appropriate functional specification for the regional case must be redesigned to help determine the size of the regional dataset market. In addition, including other aspects of the dataset (such as the level of data accuracy) in the model is a promising way of refining our analysis; including other demographic variables or buyer firms' individual conditions in the model may also generate additional insights. Applying the implications of our results to the whole economy requires us to deal with the possibility of sample selection bias. We will therefore need to select a representative sample. Expanding the scope of the target industry or incorporating other kinds of dataset in the study will enable us to draw a more holistic picture of the entire big data ecosystem. Conducting expert interviews might also help us complement our econometric outcome. Finally, if the value of location-based data is proportional to the sales volume of the user firm, estimating market size, not as a simple average figure, but as an average weighted by the turnover might produce more plausible figures [17].

[17] This was suggested by one of our reviewers.

References

Boston Consulting Group (BCG) (2012): "The Value of Our Digital Identity". http://www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf

CEBR - Centre for Economics and Business Research Ltd (2012): "Data Equity: Unlocking the Value of Big Data," Report for SAS. http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf

CHEN, H., CHIANG, R. H., & STOREY, V. C. (2012): "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Quarterly*, 36(4), 1165-1188.

General Electric (2013): "Industrial Internet: A European Perspective – Pushing the Boundaries of Minds and Machines". http://www.ge.com/europe/downloads/IndustrialInternet_AEuropeanPerspective.pdf

JIPDEC - Japan Institute for Promotion of Digital Economy and Community:
- (2012): "Report on the Study of Utilization of Personal Information by Using Anonymizing Technology," (in Japanese). http://www.meti.go.jp/meti_lib/report/2012fy/E002579.pdf
- (2014): "FY2013 Report on the Assessment of the Economic Impact from Big Data Utilization for the Infrastructure Development relating to the Trend toward the Information and Service Economies in Japan," (in Japanese). http://www.meti.go.jp/meti_lib/report/2014fy/E004190.pdf

JITSUZUMI, T. (2005): IT Investment and Its Performance in Japan: Management Strategies and Policy Initiatives (in Japanese), Fukuoka, Japan: Kyushu University Press.

LaVALLE, S., LESSER, E., SHOCKLEY, R., HOPKINS, M. S., & KRUSCHWITZ, N. (2011): "Big Data, Analytics and the Path from Insights to Value," *MIT Sloan Management Review*, 52(2), 21-31.

LOUVIERE, J. J., HENSHER, D. A., & SWAIT, J. D. (2000): Stated Choice Methods: Analysis and Applications, New York: Cambridge University Press.

MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & BYERS, A. H. (2011): "Big Data: The Next Frontier for Innovation, Competition, and Productivity". http://www.mckinsey.com/~/media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx

MIC - Ministry of Internal Affairs and Communications (2014): "Official Announcement of Quarterly Data on the Number of Telecommunications Service Subscriptions and Market Share (FY2014 Q1 - End of June 2014)," (in Japanese). http://www.soumu.go.jp/menu_news/s-news/01kiban04_02000084.html

RUSSOM, P. (2011): "Big Data Analytics," TWDI Best Practice Report. http://tdwi.org/~/media/0C630BCFD9064A9287148F1FA33460E4.pdf