

2018年6月

労働市場の民間ビッグデータ： 経済統計としての活用可能性

高田 悠矢 *

別所 英実 **

五十嵐 盛仁 ***

* 株式会社リクルートキャリア 経営統括室

** 総務省 統計委員会担当室 室長補佐

*** 総務省 統計委員会担当室 室長補佐

総務省 統計委員会担当室

〒162-8668 東京都新宿区若松町19番1号

総務省統計委員会担当室ワーキングペーパーは、統計委員会担当室スタッフ又はスタッフと外部研究者との共同による調査・研究の成果をまとめたもので、公的統計の整備に係る各種施策に役立てることを企図としている。ただし、ワーキングペーパーの内容や意見は、執筆者個人に属し、総務省の公式見解を示すものではない。

労働市場の民間ビッグデータ： 経済統計としての活用可能性

高田悠矢* 別所英実** 五十嵐盛仁***

要 旨

近年、各国の統計作成当局は、民間ビッグデータを統計作成の新たなデータ源と捉えて、その利活用に取り組み始めている。公的統計が対象とする統計分野は多岐に亘るが、労働市場の動向をより詳細且つタイムリーに把握するために、ビッグデータを活用した新たな経済指標の作成に向けた期待が高まっている。本稿では、諸外国における先行事例をサーベイし、活用事例のあるデータ源及び取組の意義を確認した。また、我が国における取組の一例として、株式会社リクルートキャリアが公表する「転職時の賃金変動状況」を取り上げて経済統計としての意義を確認するとともに、さらなる利活用の可能性を掘り下げて評価を試みた。我が国においても、こうした活用可能性の大きい民間ビッグデータについて、試行的に検討を開始することで、ビッグデータ活用を巡る先行事例を蓄積することが重要である。

キーワード：ビッグデータ、公的統計、外部労働市場、民間職業紹介所

J E L分類：C55、J01

* 株式会社リクルートキャリア 経営統括室

** 総務省 統計委員会担当室 室長補佐

*** 総務省 統計委員会担当室 室長補佐

本稿は、2018年6月27日に総務省において開催する「ビッグデータ等の利活用推進に関する産官学協議のための連携会議」（第2回会合）における報告のために行った調査・研究をベースに作成したものである。本稿の作成にあたっては、浅野和之、佐藤学、三好譲二（以上、株式会社リクルートキャリア）、上田聖、櫻川幸恵、肥後雅博、松井伸司、八木雅彦（以上、総務省）の各氏から有益な助言を頂いた。ここに記して感謝したい。ただし、本稿の内容と意見は筆者ら個人に属し、株式会社リクルートキャリアあるいは総務省の公式見解を示すものではない。また、ありうべき誤りはすべて筆者ら個人に属する。

1. はじめに

近年、政府が取り組む抜本的な統計改革では、民間企業等が保有するビッグデータを公的統計の作成に積極的に活用することで、統計調査における報告者の負担軽減や、正確で効率的な統計作成を実現することが期待されている。しかし、統計調査が、統計作成を目的として予め調査設定されたデータであるのに対して、ビッグデータの場合には一般の経済社会活動の副産物として生成・蓄積されたものである¹ため、通常の公的統計の作成プロセスとは異なるアプローチが必要となる。また、データ毎の特徴や利活用上の課題に応じたデータマネジメントが求められる²。

民間ビッグデータの活用が期待される統計分野は多岐に亘るが、経済分析への応用を企図した取組も活発化している³。我が国においては、マクロ消費動向を把握するために POS データやクレジットカード情報の活用に向けた検討が始まっているほか、物価指数作成のためにウェブ上の物価情報の活用を検討する取組も存在する。海外の先行事例や研究では、「労働市場における民間ビッグデータ」の活用事例も蓄積しつつあり、我が国においても具体的な利活用に向けた検討に向けて期待が高まっている。

こうした中、総務省が開催する「ビッグデータ等の利活用推進に関する産官学協議のための連携会議」の第 2 回会合において、労働分野におけるビッグデータの活用可能性を検討することが予定されている。本稿は、同連携会議における報告のために行った調査・研究の成果である。

具体的には、第 2 節において、労働市場における民間ビッグデータを活用した諸外国の先行事例をサーベイする。続く第 3 節では、我が国における活用事例の一つとして株式会社リクルートキャリアが公表する「転職時の賃金変動状況」を取り上げつつ、更なる活用方法の検討を通じて、公的統計よりも詳細且つタイムリーな動向把握の可能性を確認する。第 4 節では、今後の検討課題や展望について若干言及して結びにかえる。

¹ Hammer *et al.*[2017]

² 国内外のビッグデータ活用に係る取組状況や検討課題についての概要は、別所[2018]を参照されたい。

³ 岡崎・敦賀[2015]や Bean[2016]など。

2. 諸外国の取組状況

本節では、民間ビッグデータを活用することで労働市場の動向把握を試みる諸外国の取組について、「データ源」の観点から整理しつつ、主要な先行事例を紹介する。

2.1. 概観

近年、各国の統計作成当局や国際機関では、ビッグデータを統計作成の新たなデータ源として活用するために、さまざまなプロジェクトを実施している。各取組が対象とする統計分野⁴は、物価や交通、観光、不動産など多岐に亘っているが、労働市場における雇用や賃金動向の捕捉を目的とするものも少なくない。

例えば、Eurostat や各国統計局からなる欧州統計システムが 2016 年に立ち上げたプロジェクト（ESSnet Big Data）では、個別のデータ源に着目した 7 つのワークパッケージを設けており、「ウェブ上の求人情報のスクレイピング」がテーマの一つとなっている⁵。また、Eurostat と欧州委員会が 2017 年 3 月に開催したビッグデータの活用に係るハッカソン⁶では、「地域レベルにおける雇用のミスマッチ解消に資するビッグデータの活用法」が政策課題として設定されるなど、労働市場の動向把握を目的としたビッグデータ活用が統計作成当局にとって、重大な関心事項であることが窺われる。

なお、我が国においては、労働市場を巡る民間統計の公的な活用は限定的であるほか、公的統計の作成のために労働市場の民間ビッグデータの活用を企図した取組はみられていないものの、その活用可能性への関心及び期待が高まっている（詳細は補論 1 を参照）。

2.2. データ源に基づく整理

以下、諸外国の先行事例や研究事例について、統計的利活用の「データ源」の観点から整理しながら具体的な取組を紹介する。ここでは、IMF スタッフによるビッグデータ分類⁷（図表 1）に基づいて整理す

⁴ 国連統計委員会が 2014 年に設置したビッグデータに係るグローバル・ワーキンググループは、各国の取組リスト (<https://unstats.un.org/bigdata/inventory/>) をホームページで掲載しており、21 の統計分野に基づいて整理・分類している。

⁵ Swier[2016]

⁶ <http://ec.europa.eu/eurostat/web/products-eurostat-news/-/CDN-20170317-1>

⁷ 国連欧州経済委員会（UNECE）が、ビッグデータの生成源に着目して 2013 年に提示し

ると、「ソーシャル・ネットワーク情報」と「商業取引データ」のビッグデータ活用が確認される一方、「IoT 活用によるデータ」についての活用実績はみられていない。

サブカテゴリについてみると、「ソーシャル・ネットワーク情報」関連では SNS 情報とインターネット検索が、「商業取引データ」については、商取引記録、銀行決済記録、民間企業ウェブサイトの活用実績が確認される状況である。

図表 1：IMF スタッフによるビッグデータ分類

1. ソーシャル・ネットワーク情報（個人生成データ）
1100. SNS 情報 (Facebook、Twitter、LinkedIn など)
1200. ブログ・コメント情報
1600. インターネット検索 (検索エンジン<Google>)
1700. 携帯電話テキストメッセージ、通話記録、データ通信記録、 位置情報アップデート ラジオカバレッジ・アップデート オンラインニュース
2. 商業取引データ（取引仲介データ）
21. 公的機関によるデータ 行政記録情報
22. 商業取引によるデータ
2210. 商取引記録
2220. 銀行決済 ・株式売買記録
2230. e コマース
2240. クレジットカード
民間企業ウェブサイト スキャンデータ
3. IoT 活用によるデータ（デバイス生成データ）
31. センサー情報
311. 固定センサー
3111. ホームオートメーション
3112. 天候・公害センサー
3113. 交通情報・リアルタイムカメラ
3114. 科学センサー情報
312. モバイルセンサー（トラッキング情報）
3121. 携帯位置情報（GPS）
3122. 自動車走行データ
3123. 衛星画像

(注) 太字・下線は、筆者。

(資料) Hammer *et al.* [2017]をもとに作成。

たリストをもとに、IMF スタッフが 2017 年に改訂したもの。詳細は、Hammer *et al.* [2017] 9～10 頁を参照のこと。

2.2.1. 事例研究：「SNS 情報」の活用

SNS 情報については、(ビジネス特化型 SNS の) LinkedIn や Twitter を活用する事例がみられている。

LinkedIn については、2012 年に米国大統領経済諮問委員会が行った LinkedIn 社サイエンスチームとの共同分析⁸が存在する。同取組では、米国内の 5 千万人以上の登録者情報をもとに、リーマン危機前後(2007~11 年)の産業セクター別及び職種別の雇用者数の増減を捕捉している。米国大統領経済諮問委員会は、LinkedIn 上の登録情報を活用することで産業セクターや職種だけでなく、登録者の勤務先や肩書き、職歴など、公的統計以上に細かいデータを確認できる点を紹介しつつ、企業が求める労働者のスキルニーズの変化を捉えることが可能であると指摘している⁹。

Twitter については、ミシガン大学の研究者が中心となって失業に関するツイートをもとに作成した失業インデックス(The University of Michigan Social Media Job Loss Index)が存在する。同インデックスは米新規失業保険申請件数の早期予測にも活用され、予測値がホームページ¹⁰で公表されていた(現在は、公表を一時停止中)。Antenucci *et al.*[2014]では、同インデックスは公的統計の予測に活用できるだけでなく、自然災害や政府閉鎖などの特定のイベントに係る影響分析にも有効である点を指摘している。

2.2.2. 事例研究：「インターネット検索」の活用

Google 検索件数から失業率のナウ・キャストイングを試みる取組がある。一例として、イタリア中央銀行スタッフによる取組¹¹では、検索ワードを”jobs”と設定した上で、同検索件数と米失業率との相関について分析を行っている。同分析では、失業率の先行指標として注目される新規失業保険申請件数よりも高い相関性を確認したほか、フィラデルフィア連銀が実施するエコノミスト予測(SPF)をアウト

⁸ Council of Economic Advisors[2012] 187-188 頁

⁹ 同様の指摘は Bean[2016]でもみられている。具体的には、労働者のスキルや地域情報がリアルタイムに把握できる点や、時代の変化とともに登場しつつある産業セクターや肩書きを特定できる点などがメリットとして指摘されている。

¹⁰ 2014 年半ば以降、予測値と実績値の乖離幅が大きくなったことから、計測モデルを見直すために 2017 年 7 月以降は予測値の公表を一時停止している。詳細は、<http://econprediction.eecs.umich.edu/> を参照のこと。

¹¹ D'Amuri & Marcucci[2012]

パフォーマンスする結果を得ている。なお、検索ワードの決定に際しては検索件数の多寡を考慮したほか、“Steve Jobs”のように職探しと無関係なキーワードを取り除くなどの工夫を行っている¹²。

上記の研究以外にも、同様の手法で失業率を予測する取組¹³は複数みられるほか、米失業保険新規申請件数など失業率以外の動向を早期に推計する試み¹⁴も存在している。

2.2.3. 事例研究：「商取引記録」の活用

民間の給与計算アウトソーシング会社である Automatic Data Processing (ADP) 社が公表する米 ADP 雇用統計¹⁵は、米労働省労働統計局が作成する雇用統計の 2 営業日前に公表されることから、雇用統計の先行指標として注目されている。

同社は、米国における民間就業者の約 2 割の給与計算業務を実施している。連邦準備制度理事会スタッフによる取組¹⁶では、同社の保有するビッグデータを活用して雇用統計を予測するとともに、よりタイムリーに動向を捉えるための週次指標を開発している。ADP 社が保有するマイクロデータを活用することで、雇用統計の調査期間（毎月 12 日を含む週）以外における動向の把握や、給与支払い頻度に対する考察が可能となるほか、雇用統計のリアルタイムな予測だけでなく速報値に対する改定幅の予測にも有効であると分析している。

2.2.4. 事例研究：「銀行決済記録」の活用

銀行決済記録に関しては、JP Morgan Chase Institute が同行顧客 28 百万人以上の口座記録から百万人のサンプルを抽出した上で、米国における賃金のボラティリティの原因について分析¹⁷を実施した例がある。

¹² 当該取組については、欧州中央銀行 (ECB) クーレ専務理事が、2017 年 11 月に「ビッグデータによる政策分析」と題して行った講演の中でも、Google 検索を活用した先行事例として取り上げている。詳細は、Cœuré[2017]を参照のこと。

¹³ Simionescu & Zimmermann[2017]は、Google 検索件数から失業率を予測する先行研究をサーベイしている。これによると、米国のほかにイギリス、フランス、イスラエル、イタリア、ノルウェー、トルコ、ブラジルに関する研究が存在する模様。

¹⁴ Choi & Varian[2009]

¹⁵ 2006 年に月次公表を開始。

¹⁶ Cajner *et al.*[2018]

¹⁷ Farrell & Greig

同取組では、ランダム抽出した匿名化済みの口座情報を活用することで、職種や年齢、居住地域や年収、副業の状況など詳細な属性情報を踏まえた分析を実施している。結果の一例として、若年層の中でも給与水準が低い労働者において賃金ボラティリティが高く、米国西部により強い傾向を認めている¹⁸。

2.2.5. 事例研究：「民間企業ウェブサイト」の活用

民間企業がホームページ上で行う求人広告を活用することで、労働市場のタイムリーな把握や公的統計の精度向上を試みる取組が数多くみられている。求人広告を活用する利点について Mandel & Scherer[2015]は、(イ)革新的企業の殆どが求人広告を利用している、(ロ)必要なスキル情報が掲載されている、(ハ)地域情報を含んでいる、(ニ)複数の求人広告を束ねる検索エンジンの存在、(ホ)データのタイムリー性、などを指摘している。

具体的な取組事例としては、Rothwell [2012]が求人広告をもとに地域単位での雇用のミスマッチ分析を行った事例があるほか、米国カンファレンスボードが労働需要を捕捉するために求人広告数をもとに作成している求人広告指数（HWOL）¹⁹などが存在する。また、前述の ESSnet Big Data では、既存の公的統計は、職種や必要スキル、勤務地情報などに関する情報量が僅少であるとの問題意識の下、ウェブ求人情報のウェブスクレイピングに取り組んでいる。

2.3. 主要事例に対する若干の考察

これまでみたように、主な先行事例が取り扱うビッグデータのデータ源は SNS 情報、インターネット検索、商取引記録、銀行決済記録、民間企業ウェブサイトの 5 つに集約されるが、データ源ごとの取組件数には相応のバラツキがみられている。具体的には、SNS 情報、インターネット検索、民間企業ウェブサイトを活用した事例が多い一方、商取引記録、銀行決済記録を活用した先行事例は限定的である点が確認された。

¹⁸ 因みに、同取組では賃金ボラティリティの分析に加え、26万人のサンプルを抽出した上でオンライン・プラットフォーム・エコノミー拡大のインプリケーションについても分析を行っている。

¹⁹ The Conference Board Help Wanted OnLine (HWOL)は、米国における雇用需要をみる上で、労働省が公表する求人・労働異動調査 (JOLTS) と並んで注目されている。両者の相違については、Cajner & Ratner[2016]が参考となる。

これらのデータ源を活用した先行事例が僅少である背景について若干付言すると、商取引記録、銀行決済記録を巡るデータアクセスのハードルの高さが影響している可能性が高く、他のデータ源と比較した際の活用上のメリットの多寡を映じたものではない点は重要である。実際、前述の銀行決済記録を活用した事例ではデータホルダー自らによる取組であったほか、商取引記録の活用事例では、ADP社が連邦準備制度理事会スタッフに限ってデータ提供を行った事例²⁰となっており、一般に公開されたTwitter情報やネット求人広告とは異なった利用上の難しさが窺われる。

次に、これまで紹介した主な事例について、民間ビッグデータの活用上の意義を整理すると、(i) 公的統計より詳細な動向把握、(ii) 公的統計よりタイムリーな動向把握、に大別される。この点、同一のデータ源に対する取組であっても、分析方法次第ではいずれの目的にも活用できることが確認された。

従って、今後も内外の先行事例の分析方法や知見を活かしながら様々なデータ源を対象として活用に取り組むことで、公的統計よりも詳細且つタイムリーな動向把握に向けた可能性を掘り下げていくことが重要と思われる。特に、既に述べたように商取引記録、銀行決済記録については先行事例が僅少であるだけに、利用上の意義・可能性を見出すような取組の蓄積が期待される。

²⁰ 前掲 Cajner *et al.*[2018]では、表紙の脚注においてADP社に対してデータ提供に対する謝辞が述べられている (We thank ADP for access to and help with the payroll microdata that underlie the work performed within this paper.)。

3. 本邦転職市場データの活用事例

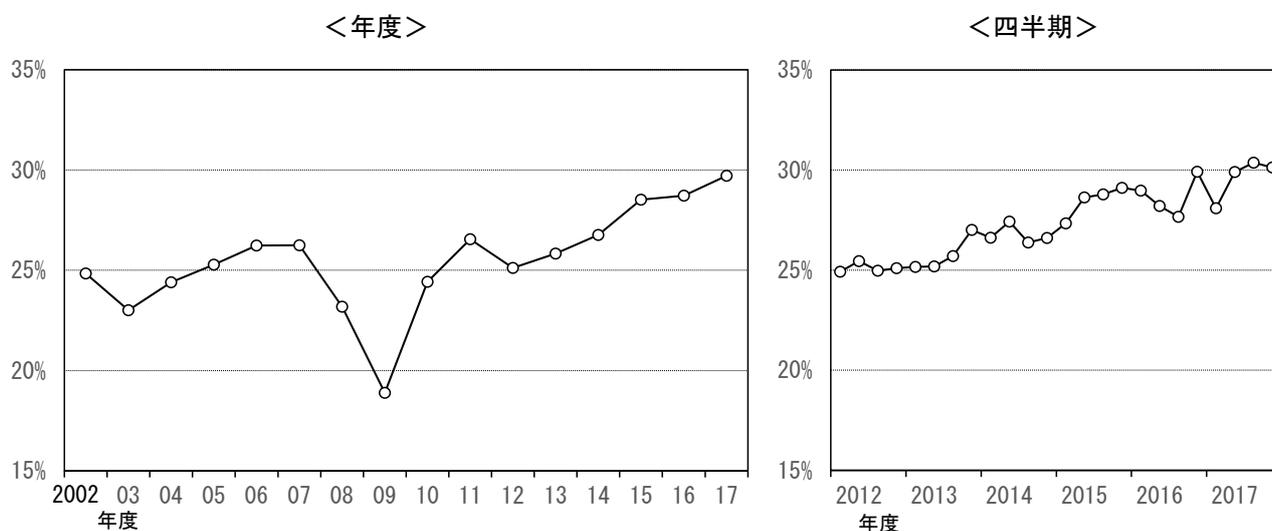
第2節で述べたように、商取引記録、銀行決済記録を活用した先行事例が僅少であることを踏まえ本節では、我が国における商取引記録の一例として人材紹介サービスを手掛ける民間企業の保有データを取り上げて、経済統計としての活用可能性を確認する。

具体的には、株式会社リクルートキャリア²¹が作成する「転職時の賃金変動状況」を取り上げ、当該民間統計の経済統計としての意義を確認する。その上で、算出上のデータ源について確認しつつ、更なる利活用の可能性を掘り下げて評価を試みる。

3.1. 転職時の賃金変動状況

株式会社リクルートキャリアでは、同社が運営する転職支援サービス「リクルートエージェント」経由の転職決定者のデータをもとに、「前職と比べ賃金²²が明確に(1割以上)増加した転職決定者の割合」を職種ごとに集計し、四半期毎に公表している(図表2、3)。

図表2：転職時に明確に(1割以上)賃金が増加した転職者の割合



(注) 算出式は図表3を参照のこと。

(資料) 株式会社リクルートキャリア

²¹ 本社：東京都千代田区、代表取締役社長：小林大三氏

²² 転職後の年収は理論年収ベース。理論年収＝月額固定給×12+賞与算定基準額×前年度実績賞与支給月数で算出される。ただし、通勤交通費、時間外・休日・深夜労働などの変動する割増賃金は月額固定給に含まない。割増賃金を固定で一律に支給する場合、一律に支給する金額は月額固定給に含む。1年未満の有期雇用契約の場合は契約期間を1年間とみなして換算した額、年俸制の場合は年俸額。

図表 3：「転職時に明確に（1割以上）賃金が増加した転職者の割合」の算出式

$$\frac{\text{「前職と比べ賃金が1割以上増加した転職決定者数」}}{\text{「転職決定者数」}} \times 100 \text{ (単位：\%)}$$

（資料）株式会社リクルートキャリア

この「転職時に賃金が増加した者の割合」は、外部労働市場における競争的な価格メカニズムの調整の結果である。従って、採用を考える企業経営者にとっては、「新規に人材を獲得する際や従業員の離職を防止するために、賃上げせざるを得ない圧力」の程度を測る上でのシグナルとして活用できる。リクルートキャリア社が上記統計の取り纏めを開始した当初の狙いはここにあった。

しかし、同統計の民間統計としての意義についてさらに考えてみると、民間企業の採用活動上の手掛かりとして活用されるのみならず、経済政策を担う政府や中央銀行などがマクロ経済の動向を把握する際にも判断材料の一つとして利用できる可能性がある。例えば、先行きの一般労働者²³の賃金動向を予測する場合には春闘の結果なども注目されるが、企業内労働市場の動向だけでなく、外部労働市場における賃上げ圧力を示す指標が存在することは有益だと思われる。

我が国においては、終身雇用・年功序列制度を採用する企業の割合が緩やかに低下しているほか、少子高齢化を背景とした構造的な人手不足も重なり、賃金決定要因において労働需給の重要度が増すシナリオも考え得る。このようなマクロ経済の潮流を踏まえても、外部労働市場をタイムリーに捉える経済統計が存在することは重要な意義を持つものと思料される。

この点、公的統計においては、厚生労働省が公表する「雇用動向調査」が、主要産業における入職・離職の状況について調査・集計している。この統計では、転職入職者²⁴の賃金変動状況についても動向を把握することが可能であるほか、多様な調査項目について精度の高い調査結果を提供しているため、経済統計としても官民さまざまな目的に活用されている。

²³ 厚生労働省「賃金構造基本統計調査」等で使用されている用語で、労働者全体からパートアルバイト等の「短時間労働者」を除いた労働者を指す。

²⁴ 入職者のうち、入職前1年間に就業経験のある者をいう。

従って、リクルートキャリア社などの民間企業が保有するビッグデータで公的統計の代替を図ることは容易ではないが、ビッグデータが有する多様性や速報性といった特徴を踏まえて民間統計を作成することができれば、公的統計よりも詳細且つタイムリーな情報を提供できる可能性がある。こうした観点から、雇用動向調査との相違点を念頭に置きながら、リクルートキャリア社が保有するビッグデータを活用した新たな経済統計の作成に向けた可能性を探りたい。

3.2. 公的統計より詳細な動向把握

リクルートキャリア社が保有するビッグデータについて、まず、職種の観点から雇用動向調査との違いを確認したい。雇用動向調査では、調査客体が図表4に示した職種分類²⁵から1つを選択して回答を行うこととなっている。

図表4：雇用労働調査における職種分類

01	02	03	04
管理的 職業従事者	専門的・技術的 職業従事者	事務従事者	販売従事者
05	06	07	08
サービス 職業従事者	保安職業従事者	生産工程従事者	輸送・機械 運転従事者
09	10	11	
建設・採掘 従事者	運搬・清掃・ 包装等従事者	その他の職業（農 林漁業等）従事者	

（資料）雇用動向調査（入職者票 様式3号）をもとに作成。

職種分類の特徴として、「専門的・技術的職業従事者」や「生産工程従事者」といったような大枠の括りで職業上の把握が可能となっている反面、例えば、自動運転技術を開発するためのITエンジニアと、企業間契約を管理する法務担当者はいずれも「専門的・技術的職業従事者」に分類されている可能性がある。

そこで、仮に統計ユーザーがこれらの職種について、それぞれ分割して動向を把握したいと考えた場合に、民間ビッグデータを使って公的統計を補完することはできないか考えてみたい。

²⁵ ここでは雇用動向調査における「職業区分」を職種として扱う。

ここで、リクルートキャリア社の保有データは、求人票に記載された業務内容を含んでいるため、これをベースとして分析目的に応じて職種分類を柔軟に設けて、整理することが可能である（同社が保有するデータの概要は補論2を参照のこと）。

現状、同社では、通常業務や市場分析の便宜上、図表5で示した職種を設定して求人票を整理している。これをもとに、一般に公表する「転職時の賃金変動状況」では、①IT系エンジニア、②機械・電気・化学エンジニア、③事務系専門職、④営業職、⑤接客・販売・店長・コールセンター、の5つの大分類について統計を作成している。

同社が保有するデータ数は、すべての職種を含めて年間3~4万件程度となっている。複数の小分類を纏めるかたちで大分類を設けて統計を作成している理由としては、トレンド把握のためには、相応の量の個票を確保する必要があるためである。従って、同社の保有するデータを活用する場合には、アドホックなかたちで職種分類を設けて分析を実施することが可能であるが、一定の統計精度を保つためには「データ量」への視点を踏まえて取り組む必要がある点には、留意が必要である。

図表5：リクルートキャリア社の職種分類

	SE		法務・知財
①	インターネット専門職(Webエンジニア含む)		内部監査・内部統制
	組込・制御ソフトウェア開発エンジニア		総務・広報
	電気エンジニア	③	人事
②	機械エンジニア		経理・財務
	化学エンジニア		マーケティング
	メディカルエンジニア		経営企画・事業企画・業務企画
	医療技術者		資材購買・物流・貿易
	食品エンジニア	④	営業
	建設エンジニア	⑤	接客・販売・店長・コールセンター
	品質管理・品質保証/認証・品質分析		流通専門職
	生産管理・生産管理コンサルタント		不動産専門職
	デザイナー		金融専門職
	制作・編集・ライター		オフィスワーク事務職

(注) ①~⑤は職種の大分類を示しており、それぞれ、①IT系エンジニア、②機械・電気・化学エンジニア、③事務系専門職、④営業職、⑤接客・販売・店長・コールセンターとして整理している。

(資料) 株式会社リクルートキャリア

3.2.1. 事例研究：IT系エンジニアに注目した賃金変動の要因分析

特定の職種に着目した分析事例の一例として、ここでは「IT系エンジニア²⁶」を取り上げて、賃金動向及び要因分析を試みる。

図表6は、リクルートキャリア社が保有するビッグデータからIT系エンジニアのみの個票を取り出して、「転職時に1割以上賃金が増加した転職決定者の割合」を作成したものである。ここでは、その経年変化について、「リーマン危機後の回復局面（2009年度～2011年度）」と「第二次安倍内閣発足後の回復局面（2013年度～2016年度）」に分けて比較を試みる。

具体的には、IT系エンジニアの「転職時に1割以上賃金が増加した転職決定者の割合」の各局面における「増加分」について、「業界」および「年齢」に着目して寄与度を求めて、両局面における要因分析を行っている。

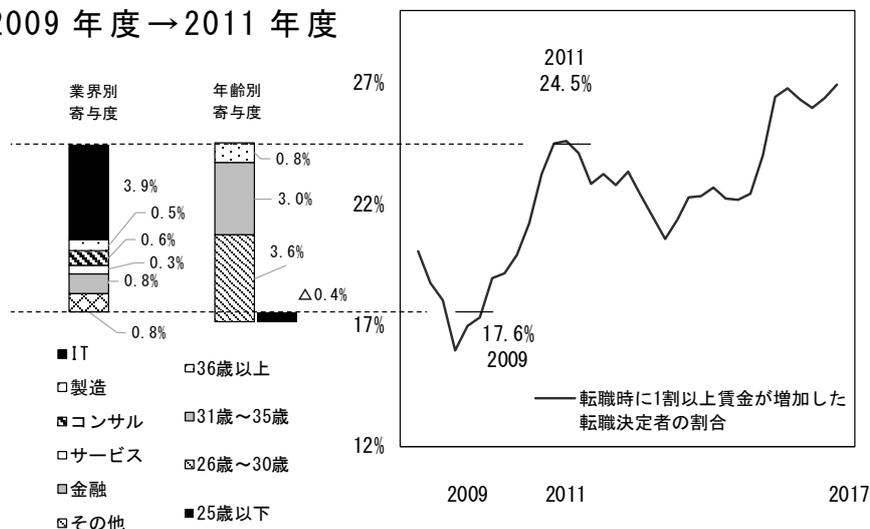
2013年度以降の局面では、リーマン危機以降の局面とは異なり、IT系エンジニアに対してより幅広い業種の企業から、さらに、従来の若手中心ではなく、より幅広い年齢層に対してより高い賃金で雇用したいとのニーズが高まったことが明らかになる。これについては、企業のIT投資が力強く持続的に増加しているという景気循環的な影響のほか、労働市場の全般的な人手不足、とりわけ、専門的な分野における人材不足の強まりといった、より長期的な要因が寄与している可能性があるとも考えられる。

以上のように、リクルートキャリア社のデータは、より詳細な職業分類で賃金変化や転職者数を把握できることから、セグメントされた専門的な人材に関する労働市場の需給構造を詳細に把握できることを示しており、日本の労働市場の分析に新たな情報を提供できる可能性を示していると考えられる。

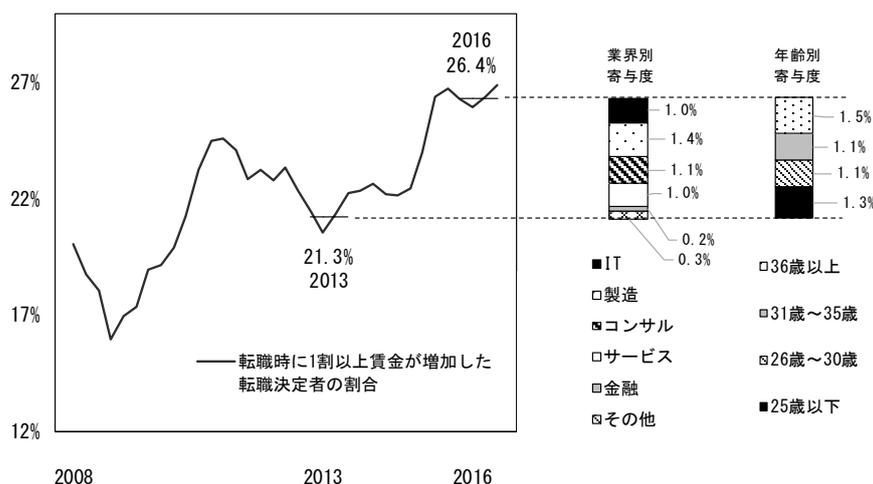
²⁶ 図表5における職種分類上の「SE」、「インターネット専門職(Webエンジニアを含む)」、「組込・制御ソフトウェア開発エンジニア」を指す。

図表 6 : 「転職時に 1 割以上賃金が増加した転職決定者の割合」の増加分に占める業界別・年齢別寄与度

▽ 2009 年度→2011 年度



▽ 2013 年度→2016 年度



▽ 寄与度分解の方法

$$U_t - U_{t-1} = \left(\frac{S_t}{N_t} \times \frac{S_{u_t}}{S_t} - \frac{S_{t-1}}{N_{t-1}} \times \frac{S_{u_{t-1}}}{S_{t-1}} \right) + \left(\frac{O_t}{N_t} \times \frac{O_{u_t}}{O_t} - \frac{O_{t-1}}{N_{t-1}} \times \frac{O_{u_{t-1}}}{O_{t-1}} \right)$$

右辺第 1 項と第 2 項は、属性 S の寄与度、第 3 項と第 4 項は、属性 S 以外の寄与度。

N_t : t 年度の転職決定者数

S_t : t 年度の転職決定者のうち、属性 S の人数

O_t : t 年度の転職決定者のうち、属性 S 以外的人数

N_{u_t} : t 年度の転職決定者のうち、転職前後で賃金が 1 割以上増加した人数

S_{u_t} : t 年度の転職決定者のうち、転職前後で賃金が 1 割以上増加した属性 S の人数

O_{u_t} : t 年度の転職決定者のうち、転職前後で賃金が 1 割以上増加した属性 S 以外的人数

$U_t = \frac{N_{u_t}}{N_t}$: t 年度の賃金が 1 割以上増加した転職決定者の割合

(注) 図は、いずれも中央 3 四半期移動平均値。

(資料) 株式会社リクルートキャリア

3.2.2. さらなる詳細性を探って

「細かい職種分類」のほかにも、リクルートキャリア社が保有するデータには、転職者の前職に関する情報も含まれることから、特定の賃金水準やスキルを持った労働者に限定して転職前後での賃金の上下を確認するなど、分析ニーズに応じた統計を作成することで経済分析上の有益なインプリケーションを引き出すことも期待される。

例えば、雇用動向調査では前職との賃金水準の変化について図表7で示した7つの選択肢から1つを選んで回答することとなっているが、リクルートキャリア社が保有するデータは、賃金の数値そのものを捕捉しているため統計ニーズに応じた区分を設定の上で統計作成が可能である。加えて、第2節で紹介したLinkedInの事例のように、企業が求めるスキルニーズの変化を把握することや、統計作成当局が時代の変化に対応するために職業分類を見直す際の手掛かりとして活用することなども期待できる。

図表7：雇用労働調査における賃金変動に関する調査項目

問 17 直前の勤め先と比べて賃金はどう変わりましたか。						
増加			変わ ら ない	減少		
3割以上 増加	1割以上 3割未満 増加	1割未満 増加		1割未満 減少	1割以上 3割未満 減少	3割以上 減少
1	2	3	4	5	6	7

(資料) 雇用動向調査 (入職者票 様式3号) をもとに作成。

3.3. 公的統計よりタイムリーな動向把握

次に、民間ビッグデータの速報性について考えてみたい。例えば、雇用動向調査は、半期統計であるために年に2度の統計調査が実施され、統計表についても半期に一度公表されている。集められた調査票を統計表として作成するためには、相応の業務が発生するため、調査実施と統計表の公表までのタイムラグは不可避となっている。一例として平成29年上半期調査を取り上げると、1月～6月に入職・離職した労働者の動向について、6月～8月に統計調査を実施して調査票を回収し、成果物たる統計表が12月20日に公表されている。

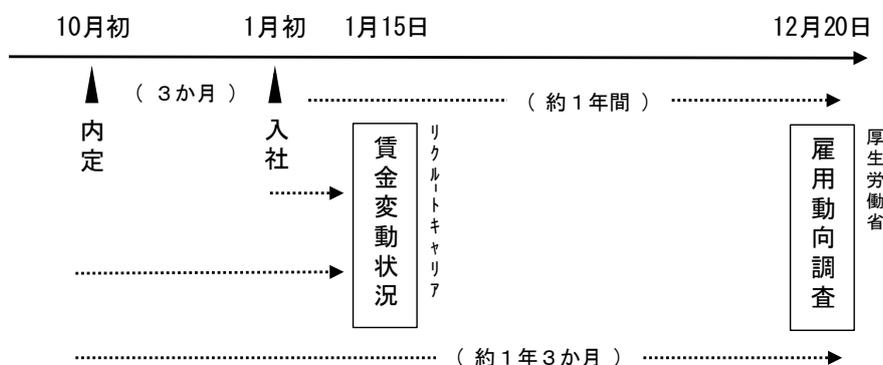
この点、民間ビッグデータを活用する場合には、業務で集めた情報を加工する時間を除けば、ニーズに応じて必要な期間を設定した上で、殆どリアルタイムに指数を作成することが可能である。現状、リクルートキャリア社は転職時の賃金変動状況を四半期単位で取り纏めているが、週次や月次で統計を作成することも理論上可能である。これは、統計調査との比較で大きな強みと言える。

加えて、雇用動向調査の場合は、調査客体が「転職後」に調査事項に回答を行うのに対して、リクルートキャリア社の統計の場合には、転職者の「内定時点」の成約賃金が捕捉されている。換言すれば、「転職前」に決定された転職後賃金（実際の入社日以降の賃金）を示すデータであるため、賃金動向の先行きを占う上で、よりタイムリーなデータを提供していると評価できる。

この点について、簡単な例を用いて再度確認しておく。雇用動向調査もリクルートキャリア社の統計も、多数の個人情報を集計の上で統計を作成して公表しているが、ここでは、ある一人の転職者（1月初に入社するケースを念頭に置く）の賃金に着目して、いつ公表される統計表に含まれるかを考える。

下図のように、雇用動向調査の場合には、入社日から数えて最大で約1年間のタイムラグが存在することとなる。ここで、当該転職者が実際に入社日（1月初）の3か月前（10月初）に内定していたと仮定して、リクルートキャリア社の統計の場合（現行の四半期公表を前提とする）についてみると、転職先からの内定から約3か月、入社日からは約2週間で結果が公表されることとなる。なお、内定時点から雇用動向調査の公表までのタイムラグは、約1年3か月となる。

図表8：公表までのタイムラグ（イメージ）



(注) 統計公表日はいずれも2017年ベース。

このように、いつの段階で入社した転職者を想定するかで異なるが、両者には相応の時間差が存在することがわかる。また、外部労働市場における価格調整が、転職希望者の内定の段階で決定される実態を踏まえると、リクルートキャリアの統計が、雇用動向調査の先行指標として活用できる可能性が確認される。

もっとも、リクルートキャリア社が捕捉している転職前賃金は、時間外労働等の変動する割増賃金を含む一方で、転職後の賃金にはそれらが含まれないため、両者を比較して作成した統計には下方バイアスが存在する。実際に雇用動向調査の先行指標を作成する際には、データ特性を十分踏まえた検討が不可欠となる。

因みに、公的統計の公表までのタイムラグを補完するためのビッグデータは、民間職業紹介所が保有するデータに限らないのは事実である。例えば、第2節の海外事例でみたように、求人広告をウェブスクレイピングすることにより、募集賃金や求人スキルの情報を集計の上で分析に活用できる可能性がある。この点、民間職業紹介所が保有するビッグデータは、募集時のデータではなく成約後の賃金である点でも、価値のあるデータと評価できよう。

4. 結びにかえて

本稿では、労働市場の民間ビッグデータを活用した諸外国の先行事例についてサーベイを行い、民間ビッグデータが、公的統計より詳細且つタイムリーな労働市場の把握に貢献する可能性を確認した。

そこで、我が国における活用可能性について、リクルートキャリア社が公表する「転職時の賃金変動状況」を一例として取り上げ、外部労働市場における賃上げ圧力を巡る、新たな経済統計の作成に向けた可能性を探った。

リクルートキャリア社の保有する転職者情報は、雇用動向調査の調査票よりも詳細なデータが含まれているため、職業分類・賃金水準・労働者スキルなどについて、公的統計では捕捉できない新たな現象を把握できる可能性がある。また、統計結果の公表までのタイムラグを大幅に短縮化できる可能性や、内定者の給与を捉えることで雇用動向調査の先行指標を作成できる可能性などが確認された。

こうした意義を有するビッグデータから新たな経済統計を作成することができれば、ユーザーの多様な統計ニーズに応えることができる可能性がある。しかし、標本調査によって母集団推計を行い十分な調査精度を確保すべく予め企画・設計された公的統計と比べると、民間ビッグデータの場合にはデータバイアスが大きいいため、経済統計を作成する場合に相応の品質を保てるか否かは大きな課題となる。

例えば、リクルートキャリア社を含む「民間職業紹介所」を經由した入職者の割合は約5～8%程度（補論3）であり、入職経路によって労働者の属性が大きく異なる傾向が強い中で、他の入職経路における転職者の賃金変化や転職パターンを捕捉できないといった限界が存在する。また、民間企業1社が保有するデータで統計的に十分な精度を確保できるかについても、統計的な検証が必要となる。

こうした課題を克服しながら、民間ビッグデータが持つ詳細性と速報性のメリットを最大化するためには、データを保有する民間主体と統計作成にかかる専門的ノウハウを持つ統計作成当局の連携が望まれる。

その際には、統計作成当局による企業や顧客情報の保護が不可欠であることは言うまでもない。また、統計作成当局が民間ビッグデータにどのようにアクセスするか、既存のノウハウで対処可能か否か

など、検討すべき事項は多い。我が国におけるビッグデータ活用に向けた検討は緒に就いたばかりであり、実務上の検討プロセスは定まっていない状況にある。まずは、本稿で紹介したような活用可能性の大きい民間ビッグデータについて試行的に検討を始めることで、我が国においても先行事例を蓄積していくことが重要である。

以 上

補論 1：我が国における取組状況

現状、我が国における労働分野の民間統計の公的な活用は、全国求人情報協会が公表する「求人広告掲載件数²⁷」など内閣府が月例経済報告における主要経済指標の一つとして注目²⁸するデータもあるが、全体としては限定的な動きにとどまっている。また、消費や物価の分野においては、統計作成当局が民間ビッグデータの活用に向けた試行的な取組を始めているが、労働市場関連の取組はみられていない。

こうした中、産業界からは、民間企業が保有する「求人情報」に着目しつつ、ビッグデータの活用可能性を指摘する声も聞かれている。例えば、経済同友会が2016年に公表した経済統計改革に向けた提言書では、職業紹介事業者の保有する求人情報を候補の一つとして挙げている²⁹。

公的統計やビッグデータに関する有識者からも労働分野におけるビッグデータの活用への期待が高まっている。一例として、柳川・渡辺[2017]では、労働分野におけるビッグデータの活用可能性について、「例えば、求人サイトの賃金や求人数などのデータを使えば賃金統計などに必要な情報を収集できる。求人サイトには必要なスキルが記述されているので、どのスキルに対する需要が高いのか、そのスキルに対する価格（賃金）はどう変化しているのか、海外の水準と比べてどうなのかといったことをほぼリアルタイムで読み取ることができる。厚生労働省等の作成する現行の労働統計では不可能だ」と具体的な活用方法を提言している。

また、統計委員会における議論において西村清彦委員長は、一例として「労働」分野を取り上げつつ、各府省に対してビッグデータ活用に向けた検討を求めている³⁰。

²⁷ <https://www.zenkyukyo.or.jp>

²⁸ <http://www5.cao.go.jp/keizai3/getsurei/getsurei-index.html>

²⁹ 経済同友会[2016] 36-37頁

³⁰ 総務省統計委員会[2017]

補論 2：人材紹介のビジネスモデルと捕捉されるデータ

人材紹介事業では、通常、企業と求職者の双方³¹に対して専任のアドバイザー割り当てている。アドバイザーには、企業の求める人材要件等の整理をサポートするとともに、求職者への求人紹介や応募書類作成のサポート、各選考プロセスにおける伴走・日程調整等を行う役割が求められる。

こうした事業モデルの性質から、人材紹介事業者は、求人広告の掲載料を収益源とする求人広告事業とは異なり、求職者の人材紹介業者への登録情報に始まって、選考・内定・入社に至る一連のプロセスで蓄積されるさまざまなデータを捉えることが可能となっている。

図表：捕捉されるデータの例

求人に関する情報の例	求職者に関する情報の例	選考プロセスに関する情報の例
企業概要 従業員数 資本金 平均年齢 所在地 株式公開状況 株主 事業内容サマリ 売上高 経常利益	プロフィール 性別 生年月日 勤務先企業 居住地	書類送付日／可否 面接日／可否 条件面談日 内定日 入社日 理論年収
求人要項 仕事内容 必要能力・経験 雇用形態 試用期間	学歴・語学・資格等 学校名 学部・学科 入学／卒業年月 TOEIC／TOEFL等スコア 資格	
勤務条件 想定年収 賃金形態 通勤手当 残業手当 退職金 社会保険 寮・社宅 所定労働時間 フレックスタイム制有無 残業状況 休日 有給休暇	職務経歴 経験社数 勤務先企業 職務 役職 給与 各種手当	
選考内容 面接回数 筆記試験有無 採用予定人数	選考内容 面接回数 筆記試験有無 採用予定人数	

(資料) 株式会社リクルートキャリア

³¹ 一人の担当が兼任するケースもある。

補論 3：転職市場における「民間職業紹介所」のシェア

株式会社リクルートキャリアが公表する「転職時の賃金変動状況」は、同社が運営する人材紹介サービス「リクルートエージェント」経由の転職者情報を用いて算出されており、当該事業は雇用動向調査における分類では「民間職業紹介所」に該当する。ここでは、外部労働市場全体における「民間職業紹介所」のカバレッジを紹介する。

雇用動向調査により、我が国の入職者全体の人数を確認すると、2016年における既就業者（転職入職者）の入職者数は4,778千人であった。

これを入職経路別にみると、最も割合が高いのが「縁故」経由で1,327千人（約28%）。次いで「広告」経由が1,294千人（約27%）。「職業安定所」と「ハローワークインターネットサービス」を合算すると、1,130千人（約24%）となっている。こうした中、人材紹介事業サービスに該当する「民間職業紹介所」経由は224千人（約5%）に留まっている。

また、入職者数4,778千人を就業形態別にみると、パートアルバイト労働者が1,841千人、一般労働者が2,937千人となっている。上述の「民間職業紹介所」経由の入職者が全て一般労働者を対象としていると仮定すると、一般労働者における「民間職業紹介所」経由の割合は、約8%程度と考えられる。

参考文献

- 岡崎陽介・敦賀智裕 [2015] 「ビッグデータを用いた経済・物価分析についてー研究事例のサーベイと景気ウォッチャー調査のテキスト分析の試みー」、日本銀行調査論文、2015年3月
- 経済同友会 [2016] 「豊かさの増進に向けた経済統計改革と企業行動ー新たな指標群『GNIプラス』の提案ー」、経済同友会
- 総務省統計委員会 [2017] 「第77回基本計画部会・第9回横断的課題検討部会（合同部会）議事録」、総務省
- 別所英実 [2018] 「民間ビッグデータを統計として活用するためには、何が必要か：諸外国の取組事例の紹介と日本における課題の整理」、総務省統計委員会担当室ワーキングペーパー、2018-WP01
- 柳川範之・渡辺努 [2017] 「民間による経済統計の革新ー統計民営化の可能性も含めてー」、『統計』2017年1月号
- Antenucci, D., Cafarella, M., Levenstein, M., Ré, C. & Shapiro, M [2014]. “Using Social Media to Measure Labor Market Flows,” NBER Working Paper No.20010
- Bean, C [2016]. “Independent Review of UK Economic Statistics,” Cabinet Office and HM Treasury
- Cajner, T., Crane, L., Decker, R., Hamins-Puertolas, A., Kurz, C. & Radler, T [2018]. “Using payroll processor microdata to measure aggregate labor market activity,” Staff working papers in the Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System
- Cajner, T. & Ratner, D [2016]. “A Cautionary Note on the Help Wanted Online Data,” FEDS Notes
- Choi, H. & Varian, H [2009]. “Predicting initial claims for unemployment benefits,” Google Inc.
- Cœuré, B [2017]. “Policy analysis with big data,” Speech at the conference on “Economic and Financial Regulation in the Era of Big Data”, organized by the Banque de France, Paris, 24 November 2017

- Council of Economic Advisors [2012]. “Economic Report of the President”
- D’Amuri, F. & Marcucci, J [2012]. “The predictive power of Google searches in forecasting unemployment,” Bank of Italy Working Paper No.891
- Farrell, D. & Greig, F [2016]. “Paychecks, paydays, and the online platform economy,” JP Morgan Chase Institute
- Hammer, G., Kostroch, D., Quiros, G. & STA Internal Group [2017]. “Big Data: Potential, Challenges, and Statistical Implications,” IMF Staff Discussion Note
- Mandel, M. & Scherer, J [2015]. “A Low-Cost and Flexible Approach for Tracking Jobs and Economic Activity Related to Innovative Technologies,” Nesta Working Paper 15/11
- Rothwell, J [2010]. “Education, Job Openings, and Unemployment in Metropolitan America,” Brookings Institution
- Simionescu, M. & Zimmermann, K [2017]. “Big Data and Unemployment Analysis,” GLO Discussion Paper, No. 81, Global Labor Organization, Maastricht
- Swier, N [2016]. “Webscraping for Job Vacancy Statistics,” Eurostat Social Statistics Conference 2016