

不動産市場分析と不動産 ID

一橋大学大学院 ソーシャル・データサイエンス研究科 教授 清水 千弘
しみず ちひろ

1. はじめに

近年、不動産テックまたはPropTechなどという言葉がしばしば聞かれるようになった。AIなどの新しいテクノロジーを使って産業そのものを根底から変革させることが期待されているようだが、現段階では、不動産価格や家賃を予測するサービスが提供されたり、他の業界でも利用されているチャットボットなどが導入されたりはしているものの、本格的なテクノロジーの導入は、まだまだこれからではないかと考えている。

このような産業界での変革は、研究開発があって初めて可能となる。前述の不動産価格の予測システムにおいても、長い研究開発期間があり、様々な学術的な論文が蓄積される中での知見が、社会で実装されるようになったに過ぎない。また、そのような研究もまた、不動産の産業界で蓄積されたデータがあって初めて可能になる。そのような意味で、現在、不動産業界で進行している新しいテクノロジーの導入は、産業界が研究者に対して情報提供といった形で貢献してきた成果を、研究者が産業界に対して、その社会実装を支援するといった形でお返ししているといったことになる。

しかし、このような不動産価格の予測モデルに関する研究をする、または研究成果をもとに社会実装するといった場合において、その根幹となる不動産価格に関するデータは、依然として大きなコストをかけなければ、実際には利用することができない。

一方、予測モデルを構築していくための頭脳部分、いわゆる広い意味でのモデルは、計量経済学的な接近法に加えて、機械学習の分野で発達した手法が利用されるようになった。20世紀後半には、「データマイニング(data mining)」または「コンピュータ経済学(Computational Economics)」と呼ばれる分野が注目されることもあったが、それらの手法が学術界でも産業界でも定着することはなかった。

古くは計量経済学的接近法では、できるだけシンプルなモデルで推計していくことが求められた。入手可能なデータや計算機の能力に制限が強かったことも、その理由の一つである。しかし、ビッグデータなどと揶揄されるように、大規模かつ多様なデータが利用可能となり、計算機の性能が飛躍的に高まる中では、機械学習の領域において開発されてきた各種手法が活躍できる環境が整備されてきた。また、計量経済学的接近法も、セミパラメトリックスモデル、非線形モデル、分位点回帰モデル、空間計量モデルなど、多くのモデルが、容易に計算できるようになった。このような計量経済学的接近法や伝統的な多変量解析、そして機械学習と、多くのモデルの利用が可能となる中では、研究対象に対して、どのようなモデルを適用していったらよいのかといったモデル選択の問題が重要になってきている。

1980年代には、機械学習における有力な手法となっていたニューラルネットワークや決定木・回

帰木などの社会実装が始められたころに、共通のデータ基盤をもとに、その予測力が競われたころがあった。そのデータが、Harrison and Rubinfeld (1978)が不動産価格の予測モデルを提案した際に利用した「ボストンの住宅価格データ」である(豊田(2001))。そのため、不動産価格の予測問題は、典型的な教材となり、代表的な統計学の教科書の一つである David ら(2023)の「Statistics for Business and Economics」でも、住宅価格の予測問題が回帰モデルの章で利用されている。また、そのデータは、30年以上前に SPSS と呼ばれる汎用的な統計ソフトウェアの演習用データとして提供されていたし、近年ではオンラインから入手することもできる。

しかし、このような手法がいくら発達したからといっても、前述のように分析に利用可能なデータがなければ、何も生み出すことはできない。不動産価格の予測モデルに関する研究を実施するためにも、予測システムを開発するためにも、不動産価格に関わるデータは必須である。データは研究の進化にも産業界の発達にも欠かすことができない社会的な資源なのである。

不動産価格には、実際に市場で取引された「取引価格・市場価格」から、不動産鑑定士といった専門家によって決定される「鑑定価格」と複数の情報源があることに注意が必要である。そうすると、どのような情報源を利用して研究を行うのか、予測システムを開発するのかという情報選択の問題は、もっとも重要な意思決定となる。その情報選択においては、価格情報だけをみて判断することはできない。不動産は、ただ一つと同じ性質をもつものがないという特性を持つ。同じ「大きさ」、同じ「形状」の土地の上に、同じ「性能」の建物が建っていたとしても、「場所」が変わればその価格・価値は変化する。同じ「場所」の同じ「大きさ」、同じ「形状」の土地の上に建物が建っていたとしても、時間が変化すれば、経年減価が起こるし、追加的な投資が行われれば価値は変化する。ここに、不動産市場分析の特性が存在しているのである。

マイクロデータを用いた市場分析における典型的な財・サービス市場では、同じ品質のものが繰り返し取引されることを前提とする。季節財のように、特定の時期にデータが欠損することは発生しても、原則としては、継続してその価格や数量のデータは利用可能である。季節財などのように特定の時期にデータが欠損したりする場合には、完全なパネル化ができなくても、アンバランスなパネルデータとして分析する(これを「temporary missing value」という)。また、パソコンや TV、車などの耐久消費財を対象としたマイクロデータを用いた市場分析においては、連続してデータを繰り返しとることができたとしても、新製品が登場することで市場から取引がなくなってしまうことで、情報が欠如する。そして、その製品は製造停止となり、二度と市場に登場することはない。このような問題は、「生産の入れ替え問題(Product Churn)」と呼ばれる。このようなデジタル財を対象とした分析では、価格・数量のデータとともに商品特性がわかるマスター情報とを紐づけながら分析を進める(Diewert and Shimizu (2024))、新井・清水(2024))¹⁾。このような分析でも、同じ製品が一定期間連続して観察が可能であることに注意が必要である。その連続して観察ができたのちは、永遠に市場に出現することはない(これを「permanent missing value」という)。

しかし、不動産市場分析では、ランダムに取引が発生し、その属性も常に変化してしまうという特性が存在している。そうすると、分析データの作成においては、このような市場の特性を踏まえてデータベースを設計しながら研究を進める必要がある。このことは、不動産データベースを設計していく上で極めて重要な性質であるということをお忘れはいけない。つまり、不動産の誕生から市場からの撤退、つまり滅失に至るまで情報を接

¹⁾ レジデータと結合されているデータは、商品コードと併せて JAN コードと呼ばれる体系で管理されることが多い。JAN コードとは、13 桁または 8 桁のバーコードで表示される商品識別番号であり。POS システムや在庫管理システムを活用して商品販売、在庫管理を行う場合には、必要不可欠なコードである。

合していくことが重要であるということを意味しているのである。

不動産価格のデータを長年にわたり生産してきたのは、不動産鑑定士、または宅地建物取引士と呼ばれる専門家である。とりわけ不動産鑑定士は、不動産価格を広い意味で予測することを生業としているので、その情報生産は、その産業の中の根幹的な役割を持つ。そして、予測問題に関して、古くから不動産価格の予測が取り上げられてきた背景には、米国でも日本でも不動産鑑定士という専門家によって価格が決定されているために「エキスパートシステム」の開発と親和性が高かったという構造に加えて、市場で実際に取引された「取引価格」が存在するために、典型的な「教師あり」データの分析として研究や開発ができるためである。

そうすると、エキスパートとして、不動産鑑定士がどのように情報生産をしているのかということが重要になる。不動産鑑定士の情報生産実務に目を向けてみると、不動産に関する属性データは、単一のデータ源から構築されるものではなく、複数のデータ源を接続することで分析用データを構築していることがわかる。登記簿情報であったり、自治体が保有する各種台帳や図面からデータを取得したりして、不動産データに情報を結合するという作業が行われている。このような複数のデータ源を接続する作業には、かなりの労力と時間、そして一定の専門性が必要とされるため、情報生産には高い費用が発生していると考えてもよい。

このような費用は、不動産市場の透明化や効率性を阻害する要因にもなっていることから、そのような社会的費用を低減させるために(国土庁(2000))、情報の接合を容易とするキイとなる ID、つまり「不動産 ID」を整備していこうとする動きが、再び活発化してきている。

ここで「再び」といったのは、不動産 ID の整備に関する議論は、筆者がかかわっただけでも、四回の研究と実践の機会があった。最初は、1990 年代初頭となる。当時は 1980 年代後半から始まる戦後最大の不動産バブルに対応するために、政策的

な対応が要請された時期である(国土庁(1992))。そのもっとも代表的な政策の一つが、国土利用計画法による取引規制であった。国土利用計画法では、「土地の投機的取引及び地価の高騰が国民生活に及ぼす弊害を除去するとともに、適正かつ合理的な土地利用の確保を図ることを目的として、『土地取引の規制に関する措置』」を定めている。最も強い規制は、規制区域を設定し、取引を許可制として実施するものである。当時は、監視区域を設定したうえで届け出制としていた。このような政策は、短期間(例えば 2 年以内)の転売を投機目的の取引として指定して、防止していくことであったため、同一取引を所有者とともに登記簿をもって特定する必要があった。当時は、不動産の取引記録を登記簿データから転記して、データベース化を全国で実施していたのである。そのための不動産 ID が必要となり、データベースの設計を登記簿データと地図基盤と結合させて、投機的取引を特定するという作業を実施していた(東京都政策報道室(1999))。

筆者がかかわった二回目の取り組みは、1990 年代半ばに実施した地価税・固定資産税との課税客体の特定である。地価税法(平成 3 年 5 月 2 日法律第 69 号)は、国税として、一定の土地等を有する個人及び法人を納税義務者として課される。バブル崩壊のタイミングで導入されたことから、地価が下落局面に入るとともに、バブル崩壊により景気後退が始まったために、地価税、固定資産税といった保有課税の負担実態を確認することが急務となった。そのような中で、企業体ごとに地価税または固定資産税の負担水準と企業の収益との関係性を把握することが要請されたが、登記データをもとに構築されている課税台帳では、その名寄せができなかったのである。例え同じ自治体内においても、所有者名で土地を名寄せすることができなかった。ましてや自治体がまたがった時に、ある企業がどこにどのような土地を所有しており、その税負担額がいくらになっているのかを捕捉することができなかったのである。そのような中で、不動産 ID を整備し、企業番号と接合することで、

企業別の負担実態を捕捉可能とした。同時に、固定資産評価実務に地理情報システム (GIS : Geographic Information System) を導入し、不動産 ID と連携させるとともに、移転登記情報と航空写真から変化情報を捉え、情報更新をしていくシステム開発を行った。のちに出てくる「法人土地統計調査」は、そのような中で始まった統計調査でもあった。

三回目の取り組みは、不動産投資市場の育成を目指し、国土審議会不動産投資検討小委員会で委員長代理を務めたときである²⁾。急速に成長する証券化市場において金融市場との融合を進めるために情報基盤を整備することが急務であった。この時期には、不動産にかかる情報が一気に多様化したのである。日本のリート市場の特性から、利益相反取引を防止するために、不動産鑑定評価が重要な役割を担うことになった。そのような中では、不動産証券化市場の不動産鑑定評価では、エンジニアリングレポートが導入され、記録しなければならない属性が増加した。さらには、半年に一度、物件ごとの収益・費用を開示することが要請されたため、その情報流通を促進させることが必要となった。この時には、不動産 ID だけでなく、EDI (電子情報交換: Electronic Data Interchange) を整備することを目指した。中央銀行や金融監督官庁がローン債権の市場整備の観点から、収支項目と費用項目の標準化を要請してきていた。そのため、不動産鑑定評価における DCF 法での情報を統一化するとともに、開示項目を統一し、それを流通させることで市場の透明化と流通コストを低下させることが実現できるものと考えていた。さらには、エンジニアリングレポートなどで生産される情報をパネルデータ化することで、プレプリントの状態での情報の更新ができるようになると、不動産鑑定評価書の生産コストも低減させることができるものという効果も想定していた。残念ながら、不動産鑑定士業界をはじめ強い反対があり実現ができなかったが、当時に実施を決定してい

れば、現在の成長した不動産証券化市場における不動産鑑定評価書の作成や取引の社会的な費用を大きく低減させることができていたものと思っている。

そして、近年では、2019 年に総務省統計委員会担当室で委員長を務めた、「不動産パネルデータベースの構築及びデータ分析に関する調査研究」³⁾ プロジェクトである。同プロジェクトは、10 年以上前に、統計審議会、その後の統計委員会で実施が決定されていた法人土地統計調査等のパネルデータ化が実現されていなかったことから、その調査設計についての実行案を示すことが統計委員長から諮問されて始まったものであった。しかし、実際の調査票の記録を見てみると、住所項目に不備があったり、5 年に一度の調査といった時間間隔が大きい調査であったりしたため、キイとできるのは座標であり、地図基盤との結合が必要であると判断した。最終的には、調査票のプレプリントの電子化を進める、という目標の下で調査を円滑に進めることができる方策を模索した。そのためには、不動産 ID を付与し、パネル化を前提とした調査設計が実現できるものと考えたのである。現在の国土交通省の議論は、この委員会の参加者が中心で進められているが、人選も含めて誤ったメッセージを出してしまったのではないかと危惧している点はある。

本来、不動産 ID と地図基盤は独立に整備を進める必要がある。もっと言えば、不動産 ID、住所、そして地図基盤はそれぞれ独立なものであり、それぞれの最適化の下で整備すべきものである。その後、3 つの情報源の接合ルールを定める方がよい。前述のように、統計調査の調査票上の課題と調査間隔が大きいということで、総務省では座標に頼った地図基盤をベースに制度設計をせざるを得なかった。しかし、不動産 ID に関わる前述の 3 つのプロジェクトは、地図基盤に頼るのではなく、地図基盤の一定の限界を前提としたうえで、

²⁾ https://www.mlit.go.jp/policy/shingikai/s104_fudousan01.html

³⁾ https://www.soumu.go.jp/main_content/000635188.pdf および https://www.soumu.go.jp/main_content/000693243.pdf

不動産 ID の整備を提案したものであり、追時的な動態的な不動産 ID として設計する必要がなかったために、地図を参考として利用することを想定した。5 年に一度という統計調査の特殊性と統計調査という一定の誤差を吸収するための次善の策としての判断であった。

このような情報生産は、不動産市場分析に限らず、実証分析を行う研究者、データ基盤を用いて新しいサービスなどを開発していく実務者双方にとって極めて重要な能力となる。ここでは、「不動産市場分析と不動産 ID」として、不動産の価格予測を行う研究を推進するという題材をもとに、情報整備の進め方と、不動産 ID に期待する点を整理したい。

2. 不動産市場分析のデータ資源⁴⁾

2.1. 不動産価格の種類

わが国において、不動産データを用いたデータ分析をしようとした場合、様々なデータの中から、その研究にふさわしいデータを選択することから始めなければならない。先行研究を見ると (Shimizu and Nishimura (2003), (2006))、不動産データには市場データと鑑定評価データが存在しており、市場価格または鑑定評価にも、複数の種類のデータが存在している。

日本の不動産価格、土地価格においては、古くは「一物四価」と揶揄されたように、政府が公表する鑑定価格だけでも複数の価格が存在していた。正確には、現在においても、その状況が変わることではない。そうすると、研究者は、そして不動産価格情報にはどのような情報があり、どのような性格を有するものであるのかを理解しておく必要がある。

わが国で公的部門により公表される地価情報としては、国土交通省による「公示地価」、各都道府県による「基準地価」、国税庁による「相続税路線価」、各市町村による「固定資産税路線価」といった 4 つのマイクロデータがある。それぞれの情報

は目的に応じた特色を持つ。相続税路線価、固定資産評価額は課税目的の価格情報であり、路線を単位として情報が提供されている。公示地価と基準地価は、それぞれ補完関係があるため地価の定義や価格水準は整合しているが、評価時点や調査対象地域が異なる。

一般的に市場で取引されている財やサービス、資産などは、「価格」といった場合、「取引価格」をさす。この「取引価格」は市場で決定され、流通しているため、「市場価格」とも呼ばれる。この市場価格にも、Shimizu et al (2016)で整理されているように、オンライン上や新聞・情報誌などで掲載されている募集価格 (Asking price) から、実際に取引された価格といった意味で、複数の価格が存在する。日本では、オンラインや情報誌に掲載された募集価格 (Asking Price)、不動産流通業者間で共有されているレインズ (Real Estate Information Network) と呼ばれるシステムに登録されている募集価格および取引価格、または国土交通省が整備している「取引価格データ」が存在する (Shimizu et al. (2016))。

取引価格、公示地価、基準地価は、国土交通省の「土地情報総合システム」からデータを閲覧またはダウンロードすることができる⁵⁾。固定資産税路線価は、一般財団法人資産評価システム研究センターから開示されている⁶⁾。

「公示地価」「基準地価」データは、電子化による開示がすすめられた結果、かつては電子化されている情報項目に不足があったが、現在では、すべての情報が整備され、公開されている⁷⁾。デー

⁵⁾ <https://www.land.mlit.go.jp/webland/> からデータを得ることができる。しばしば日本の不動産情報の整備や開示が遅れているということが指摘される。筆者の一人も四半世紀前にはそのような指摘をしてきたが、現在では、国土交通省、不動産鑑定士協会連合会の尽力により、米国・英国、オランダと並び、世界で最も情報整備と開示が進んでいる国である。

⁶⁾ <https://www.recpas.or.jp/new/teisyutsu/index.html>

⁷⁾ かつては、昭和 40 年代のデータには住所・駅名などのテキストデータが入っていないなどの問題があったが、現在ではすべてのデータが電子化され、1970 年からのすべての情報が利用できる。また、国土数値情報と

⁴⁾ 本章は、Nishimura and Shimizu (2003) を要約したものである。

タには、住所（地番および住居表示）、地価情報（当年公示地価および前年価格・対前年変動率）、敷地条件（整形地かどうかを示す形状区分・地積・間口・奥行比）、街路条件（前面道路幅員・方位・舗装状況）、基盤整備状況（水道・下水道・ガスの有無）、交通利便性（最寄駅および駅までの道路距離）、公法上の規制（都市計画用途地域・容積率・建蔽率・高度地区かどうか等々）、周辺の土地利用の状況など、地価形成に影響すると考えられる多様なマイクロな情報が記載されている。一方、路線価や取引価格、オンラインなどで収集可能な募集価格などでは、その属性は、公示地価ほどに豊富ではなく、その測定も申告であったり、専門家が生成していなかったりする場合もあることから、誤差が大きい可能性がある。

土地を含む不動産は、同一の財が存在していないという特性を持つことから、その属性によって価格が変化する。そうすると、不動産市場分析を行うためには、価格調査と併せて、不動産の属性データが重要になる。その意味で、不動産価格データといったときには、「価格」データだけでなく、その価格に付随する属性データも多様さと正確度が極めて重要であるということ認識しておかなければならない。国土交通省、不動産鑑定士協会連合会において、公的評価に関わる不動産鑑定士による調査によって、このような情報が専門家によって生産されているのである。

2.2. 鑑定価格・評価価格

研究者または市場分析を行う実務者は、分析で用いるデータの性質を正しく理解しておく必要がある。ここで、鑑定価格・評価価格を説明する。鑑定価格は、不動産鑑定士によって評価された価格を意味する。わが国における不動産鑑定評価制度は、昭和38年の「不動産鑑定評価に関する法律（昭和38年法律第152号）」に基づき確立されたものであり、費用から算定する「原価法」、土地の収益を「適正な割引率」を設定した上で現在価値と

して求める「収益還元法」、近隣の相応する土地の取引事例をもとに求める「取引事例比較法」の3手法を比較考慮した上で決定される。

同制度は、昭和38年6月8日に建設大臣から「最近における宅地価格の騰貴及び宅地の入手難が、国民経済の健全な成長及び国民生活の安定に重大な障害を及ぼしている現状にかんがみ、宅地価格の安定、宅地流通の円滑化、宅地の確保及び宅地の利用の合理化を図るために、いかなる制度上の措置を講ずるべきか」という諮問を受け、宅地制度審議会において審議が開始され、制度化にいたった。

つまり、初期の問題意識のなかには、「地価抑制」という考えが前提にあり、そのために「正常価格」という概念が登場する。そして、その「正常価格」においては、「あるべき価格(sollen)」として評価すべきか「あるがままの価格(sein)」として評価すべきかといった議論が長く続いた。そのような定義の変更を批判的にとらえる不動産経済学者もいるが、経済測定の分野においては、当然なこととして考える。それぞれの時代における地価という情報に対する需要によって測定すべき定義が異なってきたと考えるべきである。例えば、国民経済計算における国内総生産(GDP: Gross Domestic Product)の測定や消費者物価指数(CPI: Consumer Price Index)の測定においては、時代の変遷と併せて、その測定マニュアルは改訂をされてきた。調査が開始されていた時代にはなかったデジタルカメラやパソコンのようなデジタル財が登場したり、Air B&B やライドシェアのような新しいサービスが登場したりしてきているのもその一因である。土地価格においても、1970年代には公共事業が活発であり、用地取得のために全国で統一的な基準として土地価格を決定していくことが重要であった。また、不動産バブルによって経済システムに不安定性がもたらされている中では、「あるべき価格」を専門家が社会に提供することが要請された。不動産バブルが崩壊した後は、証券化市場を整備していくことが官民一体となって進められたが、金融市場との融合の中では、「あるがまま

連携したことで、従来は取得が困難であった座標も付与されている。

の価格」を付けていくことがとりわけ都市部においては要請されたことは自然なことである。しかし、社会全体が縮退していく中で、その測定基準を継続していくべきか、新しい定義をすべきかは、今後議論をしていかなければならない。GDP においては、「Beyond GDP」と呼ばれるように、成熟した社会において「well-being」の測定が重視されている。土地価格は、社会全体の「well-being」の測定において重要な役割を持つことも考えられるのである。公示地価の問題は、公的な統計調査のように、調査計画の改訂を行う手続きが不透明であったり、開示が十分でなかったりする点である。

この鑑定価格 (appraisal price) と類似する概念に、課税のための「評価価格 (assessment price)」が存在している。外形的に評価が必要とされる税の評価は、固定資産税は市町村長が、不動産取得税における評価は原則として固定資産税における評価に基づき都道府県知事が、相続税・贈与税における評価は税務署長が、登録税における評価は税務官が行うものとされている。それぞれの税の目的が鑑定評価という「正常価格」が基準となりながらも、想定する市場と異なることから、それぞれの価格が存在している。

古くは、固定資産税評価と相続税のための路線価評価においては、各自自治体及び国税庁担当者が独自に評価をしていたことから、固定資産税路線価においては自治体間や同一自治体内においても用途間などにおいて均衡が保たれておらず、さらに固定資産税路線価と相続税路線価との間にも大きな乖離が存在していた。そのため、取引価格・公示地価とあわせてしばしば一物四価と揶揄され、社会的な問題にまで発展した時期があった。そのようななかで、1989年に制定された「土地基本法」、1991年に閣議決定された「総合土地政策推進要綱」に基づき、各公的土地情報間の整合性を確保することの必要性が指摘され、相続税路線価は1992年以降公示地価の8割を目途に、固定資産税路線価は1994年評価替え以降では公示地価の7割を目途に評価が行われている。そうすると、公示地

価が重要であることが理解できよう。

地価公示は、「一般の土地の取引に対して指標を与え、及び公共の利益となる事業の用に供する土地に対する適正な保証金の額の算定等に資し、もって適正な地価の形成に寄与すること」を目的として1970年に調査が開始された。より具体的には、1) 一般の土地取引に対する指標の提供、2) 不動産鑑定士等の鑑定評価の基準、3) 公共用地取得の算定基準、4) 収用委員会の補償金額の算定基準、5) 国土利用計画法による土地取引規制における価格審査の基準、6) 国土利用計画法に基づく土地の買収価格の算定基準、としてスタートしたのである。公示されるのは、毎年1月1日における標準地の単位面積当たりの正常な価格である（法第2条第1項、規則第1条）。調査方法としては、土地鑑定委員会が2人以上の不動産鑑定士又は不動産鑑定士補の鑑定評価を求め、その結果を審理し、必要な調整を行って判定する事とされている（法第2条第1項）。

また、調査対象地域は、「総理府令で定める都市計画区域（都市計画法（昭和43年法律第100号）第4条第2項に規定する都市計画区域をいい、国土利用計画法（昭和49年法律第92号）第12項第1項の規定により指定された規制区域をのぞく）」（地価公示法（昭和44年法律第49号）第2条第1項）。

鑑定評価手法としては、取引事例比較法、収益還元法及び原価法の3手法により求められる価格を勘案して行うものとされている（法第4条）。市場の性格に応じて、戸建て住宅市場などでは基本的には取引事例比較法が中心となり、商業地などでは収益還元価格が重視される。

以上の整理からもわかるように、わが国の不動産市場分析においては、取引価格とともに、公示地価、基準地価、固定資産税評価および相続税のための路線価評価として公開される「路線価」といったデータが利用できる。ただし、そのデータの性質や発生プロセス (Data Generation Process) を正しく理解したうえで、データ分析をしていかなければならないことは言うまでもない。

さらに重要な点は、不動産に関わる情報生産には何重にも費用がかけられているということである。これらの情報生産プロセスを不動産 ID の下で集約してしまえば、不動産 ID そのものの精度も向上していく可能性が存在しているのである。

3. 不動産市場分析とデータ資源

3.1. 不動産情報の生産

現在の不動産情報の生産は、不動産鑑定士といった専門家に大きく依存し、献身的な尽力によって情報整備が実施されている。市場分析を行う専門家及び研究者は、与えられたデータをそのまま分析するのではなく、自らで情報を生産することが極めて重要な技術であり、能力となる。

不動産情報の整備では、不動産は土地と建物から構成されるため、土地・建物の性能や市場価値に対して影響を持つ属性を生産することを意味する。土地は永続的に存在するものの、建物は耐久性が長いものの有期であること、耐久性があるために二次市場が存在すること、建物は維持・修繕とともに数年に一度は大規模な修繕投資が行われることで変質していくこと、同じ敷地内に建物が連結して追加されることがあること、最終的には建物が建て替えられること、そのような時には、土地は分筆されたり、合筆されたりして面積・形状が変化すること、などといった動態的な変化がもたらされる。

市場分析では、クロスセクション分析として時間を固定化して分析をする場合と、都市・不動産市場の動学的な変化を分析することがある。データをパネルデータ化して分析する場合には、データベースの構築方法が分析する目的に応じて異なることになる⁸⁾。ある時期を出発点として分析を始めることになるが、情報の生成・生産プロセスとの関係で、土地・建物の情報について整理して

みよう。不動産は、土地の上に建物が生産される。この段階で、土地・建物に関する情報の生産と蓄積が始まる。まず土地に関しては、不動産登記によって所有者や抵当権の状況などが確定できなければ、建物の設計・計画に入ることができない。ここで、面積や公図から位置や土地形状などが確定できる。建物は計画、設計、施工を経て竣工されるが、その際に民間において多くの情報が生産され、公的部門においてもそのような情報が報告され、行政記録情報として一定期間保管されている。それらの情報を事後的に統合しようとしても、公的書類は異なる機関（中央官庁・都道府県・市町村）、または民間でも設計事務所・施工会社・建築確認申請においても「特定行政庁」または「確認検査機関」など別々に保管されていることが一般的である。それを事後的に統合することは、困難であることから、同一建物を追跡するかたちで申請書や図面を取得することができない。

例えば、建物は基本設計を経て建築確認申請が行われるが、そこで発行される確認申請番号は、必ずしも登記や他の統計に結合することはできない。建築着工統計は建築工事届に基づき付番され、全数調査のために建築される建物の戸数・面積などの数量情報と、工事費の見積額といった価格情報の詳細が理解できるが、建築確認申請と結合できないことが多いため、情報は限定的となり、建物の性能を詳しく知ることができない。また、仮にできたとしても、竣工までの間に提出される変更届の内容は反映されていない。さらに、建設工事受注動態統計調査は建設業者の受注高を抽出調査で集計しているが、元々の工事費と紐づけられておらず、工事の変更や建築価格の時間的な変動までは正確にわからないし、抽出調査であることから、すべての建物が網羅されているわけではない。

そのため、不動産取引価格情報は、不動産鑑定士によって再生産されなければならない。そのような情報がないために、不動産鑑定士は外観から判定できる情報や役所調査などから得ることができる情報だけに限定されてしまっている。また、

⁸⁾ Shimizu, Karato and Asami (2010) では、東京都区部の都市計画基礎調査を用いた建物利用の変化を分析している。データベースの構築では、同じ形状で連続的に観察できるデータ群、建物の分割が発生したデータ群、建物が統合されたデータ群を時間的に追跡できるように ID を付与して、パネルデータを構築した。

建築費については再建築原価（新規建築費）を鑑定士が計算しなければならない。また、固定資産課税当局においても、新規建築費は評点制度を用いて独自計算をしており、極めて大きな社会的費用を発生させている。

3.2. 不動産予測モデルにおけるデータ生成

情報生産だけでなく、その保管には高い費用が発生するため、各工程で清算されるすべての情報を蓄積し、保存をしていくということは現実的ではない。情報を特定し、優先順位を付けて整備・保存をしていくことが要請される。その選択には、情報の重要度と精度・正確度・費用とを見極めようで決定していくことが必要となる。

ここで、取引価格情報を用いた不動産価格の予測モデルの推計例から、必要とされる不動産の情報項目と整備状況を見てみよう。Nishimura and Shimizu (2003)では、取引価格データを取引事例と呼ばれる紙データの収集を行い、分析用データに変換したうえで、不動産価格関数を推計している。不動産価格を予測するための変数としては、「地積：LA(m²)」、「前面道路幅員：RW(cm)」、「最寄り駅」および「最寄り駅までの距離：ST(m)」、公法上の制限としては、「容積率：YK(%)」や「建蔽率(%)」などを選択している。これらの変数は、不動産鑑定士が不動産価格を決定するうえで利用してきた情報であり、取引事例と呼ばれる基礎資料に記載されている。東京都から提供されたデータを分析しようとしたときに、予測モデルの推計に重要な変数と予想される項目についても、多くの欠損値が存在していた。そうすると、一般の統計調査と同様に、欠損値を補完することが必要となる。疑義照会に該当するものが、現データに当たり、正しい数字を投入することである。そこで地積情報が未記入なものについては、1987年以降のものについてのみ登記済異動通知（土地取引電算データ⁹⁾）のデータと結合させることでデータ

取得した。このことは、不動産取引データと登記データが結合することができれば、「地積」に関しては調査が必要ないことを意味する。

続いて「前面道路幅員(10cm)」、「最寄り駅」および「最寄り駅までの距離(m)」、「容積率(%)」、「建蔽率(%)」については、未記入データの存在とともに調査員ごとの測定誤差のばらつきといった問題があった。この問題に対しては、住所データおよび付属地図データをベースとして、ゼンリン「住宅地図」および「路線データベース」をもとに電子地図上にプロットして、道路台帳、都市計画用途地域の地図など、自治体が保有するデータと結合させていくことで、「前面道路幅員」「最寄り駅」および「最寄り駅までの距離」「容積率」「建蔽率(%)」を再取得した。このような事実もまた、座標が正しく取得でき、自治体の各種台帳や図面、その他の関連地図と結合ができれば、新しい情報生産は必要ないことになる¹⁰⁾。

そのように作成されたデータベースを用いて、次のような不動産価格関数を設定し品質調整済の価格指数を推計している。

$$\begin{aligned} \log L P_{it} = & a_0 + \sum_i a_{1i} \cdot \log X_i \\ & + \sum_k a_{2k} \cdot RD_k \\ & + \sum_{i,k} a_{3ik} \cdot (\log X_i)(RD_k) \\ & + \sum_t a_{4t} \cdot TD_t + \varepsilon \end{aligned}$$

¹⁰⁾ このような一連の作業においても、結合ができないデータが一定程度存在していた。排除されたデータは、住所データ・付属地図情報から正確な地点としてプロットすることができなかった事例、登記済異動通知とリンクできなかった事例である。緯度・経度が特定できない場合には、最寄り駅の確定・駅までの距離の測定ができない。清水(1998)では、取引事例に記載されている駅までの距離の検証について指摘している。

⁹⁾ 登記済異動通知は、昭和62年以降各都道府県で電算化されていた。本データは東京都政策報道室（現・知事本部）により提供を受けた。

LP_{it} : i 種類 t 時点の地価(取引事例), ($t = 1975 \dots 1999$)

X_i : 主要説明変数

$X_1 = LA$: 地積(m^2)

$X_2 = RW$: 前面道路幅員(10cm)

$X_3 = ST$: 最寄駅までの距離(m)

$X_4 = TT$: 都心までの時間(分)

$X_5 = YK$: 指定容積率(%)

RD_k : 沿線ダミー($k = 0 \dots K$)

TD_t : 時間ダミー($t = 0 \dots T$)

専門家が作成している情報といえども、現在のデータ資源だけでは研究に利用することができないため、複数の情報源を結合して、データベースを構築しない限り、不動産価格の予測モデルを作ることにはできないのである。これは、政策や実務で利用する場合も同様となる。Nishimura and Shimizu (2003)は、土地だけを扱ったモデルであるが、建物を含む不動産では、建物階数、建物面積、高さ、建築後年数などが重要となる(Deiwert and Shimizu (2015), (2016), (2020), (2022))。さらには、不動産価格は、周辺環境などにも強く影響を受けることから、地域間の相違を考慮できるような変数を追加していかなければならないのである(Shimizu (2014))。

3.3. 不動産IDに期待すること

現在においても、不動産情報またはそれを支える地図基盤などが整備され、電子化がすすめられたことで、大量のデータが利用できる環境となった。しかし、依然として取引価格情報の中で記録されている属性データには、多くの欠損値が存在していたり、追加的に調査をしないデータが多く残されていたりする。つまり、現行の情報基盤システムでは、不動産市場分析を行うためには追加的な調査が必要となっているということは、実際の実務においても、課題を多く残していることを意味する。

不動産市場分析を行う専門家である不動産鑑定士は、独自調査をして情報生産をしているし、不動産仲介業を行う宅地建物取引士もまた、物件調査を行い、重要事項説明書などを作成している。ここにかかる時間または人的資源は、まだまだ大

きい。そのような中で、不動産IDが整備され、土地と建物の一連のライフサイクルで生成される情報が接続することができたときには、公共部門、民間部門ともに、これらの費用のみならず、大きな費用が節約できるだけでなく、統計整備をはじめ多くの領域に発展的に情報活用が可能となることが予想される。

具体的には、更地→設計→確認申請→工事届→着工→監理→竣工→表題登記→運用→維持管理・修繕→解体→滅失登記→更地といった不動産の土地建物の全体のライフサイクルごとに生産される情報を接合し、情報の資源化を進めればよい。

一つの例を挙げれば、建物竣工までの流れは、民間企業では図面作成を通して情報の生産と蓄積が行われている。しかし、現段階では紙であったりPDFであったりと、画像イメージでしか保存されていないことが多いが、BIM(Building Information Modeling)が普及してくるにつれて基本設計、詳細設計、竣工、維持管理の各段階で情報が更新され、保存されていけば、追時的に公的書類とBIMを組み合わせることで、部材レベルの単位から統一的なデータ構築ができる。このようなイベントごとに生産される情報をいかに接合させていくことが重要であり、そのためには、不動産IDは必要不可欠なものなのであると考えている。

このような情報システムの設計においては、経済統計の分野で発展してきた理論や公理、システム設計なども有用である。例えば、時間軸上で、土地が分割されたり、併合されたりすることがある。その時には、指数理論という時間反転性という公理基準からみて正しく制度設計がされているのかどうかを検証していくことが確認しなければならないことを意味する。つまり、時間軸で過去から未来にむけて集計した場合と、未来から過去に向けて集計した場合で統計が再現できるのかどうか、一致することができるのかどうかといったことを意味するものである。

統計整備のための不動産IDではなく、様々な形で不動産市場に向き合う「不動産ID」は、不動産

市場分析でいうクロスセクションとしての横断的な精度と、時間軸上でのパネルデータ化したときの精度を確認しながら設計していく必要がある、つまり、「動態的な不動産 ID」として整備しようとしたときに、どのような条件を具備させ、どのような公理の下で構築していくのかという思想設計が重要であると考えている。

4. 結論—未来型不動産情報整備システム—

情報生産技術は、実証分析を行う研究者にとって極めて重要な能力の一つである。研究に必要な情報を探索し、入手し、そのデータを分析用データに変換していく手続きは、時間がかかるだけでなく、労力も大きい。データ入手にはデータ所有者と交渉しなければならないことも多く、高い費用を支払うことも要求されることがある。そのために、競争的研究費を獲得しなければならないことも少なくない。また、金銭的な費用だけでなく、データ提供の対価としてデータの所有者との交渉の中で、データ所有企業の価値につながるような貢献が要請されて、多くの時間がとられてしまうことも少なくない。さらには、その情報が獲得できたからといってすぐに分析ができるわけではなく、複数のデータを結合させたり、自ら測定したりして追加していくことも行う。

しばしば不動産市場分析を行う研究者が、日本では情報が少なく、世界的な研究競争ができないと言うことがある。それは間違いであると断言できる。日本ほど情報が整備され、公開されている国は少ないといっても良い。欧米の研究者もまた、情報を入手するために公的部門や企業と交渉をし、複数のデータと結合させて分析用データを整備し、自らが情報生産をしている。このような努力をしても情報が得ることができない国の方が圧倒的に多い。そもそもデータが記録・保存されていないのである(国土庁(1995))。

一方、データベースの設計においては、どんなにきれいな絵を描いたとしても、実際に実行してみても初めて分かる課題もある。研究者が、分析用データを収集・構築して分析しようとして初めて

発見される課題が存在しているということと同様であろう。

本稿では、このような問題意識の下で、不動産市場分析を行うという実際の手続きの中で、情報生産がどのように行われ、その中で不動産 ID が存在するとしたら、どのような効果が期待されるのかを整理した。第三節では、不動産 ID に期待する方向性を示した。しかし、第三章で示したような情報接続と流通を実現していくためには、政府が担わなければならない課題が多く存在している。

BIM の情報蓄積、整備流通においては、統一的な情報コードが整備できていないことで、情報が生産されたとしても、個社単位で蓄積されているだけであり、情報の社会資源化ができていない。つまり、情報を流通させることができないのである。地図基盤のベースとなる住所データにおいても、総務省も選挙を担う部署と国勢調査などの統計を担う部署でも異なり、国土地理院、法務省などもまた独自の定義の下で住所を作成されている。そのため、ベースレジストリーの整備も調整が困難な状況にある。さらには、地図基盤もまた、それぞれの目的ごとに地図は存在するために、すべての住所情報と整合させることは困難である。

このような社会基盤としての課題と技術的な課題は、未来志向で制度設計、システム設計をしたときには、時間とともに解決していくものであるという姿勢が重要となる。例えば、固定資産評価システムの電子地図基盤を作成したときには、税務行政が要求する精度とは乖離が存在していたことは容易に予想できるであろう。登記簿で識別できる土地単位と実際の利用単位は異なる。固定資産評価においては、建物利用と併せて土地の一面地認定をしていかなければならない。例えば、地籍調査や登記簿によって知ることができる 100 m² の土地が 3 つあり 300 m² の一面地に建物が存在しているとしよう。その時に、100 m² の土地評価を 3 倍した評価額と、300 m² として評価した評価額は大きく変わってしまう。100 m² の土地が間口狭小で奥行きが長大だとすると、一つずつの土地価格は低く補正されるが、300 m² とまとまることでそ

のような補正の必要がなくなると、土地の価値が正常化するためである。分割されていたら無道路地として識別されてしまうようなこともある。国税レベルでは、1990年代後半になると、国内の不動産に関する情報整備後に、日本の家計、企業が海外に保有する不動産の捕捉が要請された。このような問題を解決していくためには、数十年にも及ぶ時間軸でとらえ未来志向でデータベースを設計しなければ実現はできなかつたし、その時に開始していなかったら永遠にできなかつたことになる。現在の不動産IDの設計では、まだまだ多くの問題には対応ができないが、それを未来志向でとらえれば、未来には優れた社会基盤として、日本の行政サービス、民間サービスの基盤になっていることは確実である。

税務データを公開すべきだという意見もあるが、税務情報を統計作成やその他の基盤情報として他の行政機関と共有している国は、筆者が知る限り、一部の例外を除いて存在していない。それぞれの最適化のためにデータベースは構築されているために、その目的を阻害するようなノイズはできる限り排除しないといけないというのは、データベースの構築の上での原理原則である。不動産情報において、レインズデータの公開を要求するようなこともしばしば聞かれるが、これは税務情報の利用問題と同じ構図だと考えられよう。情報生産は、それぞれに目的があり、その目的が十分に達成できていないときには、個別最適化を図ることが最善の選択なのである。

そうしたときに、現在の制約の中で、どのように不動産IDをはじめ、都市不動産市場を進化させていくための情報基盤を構築・設計していくかにおいては、現在の最適化を求めるのではなく、現在の課題を特定化し、未来志向でシステムをデザインしていくことが優先される。時間の経過の中でシステムの精度が向上し、進化していくための情報生産のフローそのものを確立していけばよいのである。近年に始まった不動産IDの整備事業は、始まったばかりである。現在は、利用可能な電子情報も官民の中で蓄積されてきていることを考え

ると、それぞれが連携していくことで、近未来においては、世界で最も不動産情報整備が進んだ国へと進化している可能性は極めて高い。

不動産IDだけでなく、EDIも含めて、再度、具体的な実務レベルまで利用可能な形で制度設計がなされ、情報整備が進められていくことを心から願うものである。

参考文献

- David R. Anderson, Dennis J. Sweeney, James J. Cochran, Jeffrey D. Camm, Jeffrey W. Ohlmann, Michael J. Fry, Thomas A. Williams, (2023), *Statistics for Business and Economics, 15th Edition*, Pearson.
- Diewert, W. E. and C. Shimizu (2015), "Residential Property Price Indexes for Tokyo," *Macroeconomic Dynamics*, 19(8), 1659-1714.
- Diewert, W. E. and C. Shimizu (2016), "Hedonic Regression Models for Tokyo Condominium Sales," *Regional Science and Urban Economics*, 60, 300-315.
- Diewert, E. and C. Shimizu (2020), "Alternative Land Price Indexes for Commercial Properties in Tokyo," *Review of Income and Wealth*, 66(4), 784-824.
- Diewert, E. and C. Shimizu (2022), "Residential Property Price Indexes: Spatial Coordinates versus Neighbourhood Dummy Variables" *Review of Income and Wealth*, 68(3), 770-796.
- Diewert, E. and C. Shimizu (2024), "Product Churn and Quality Adjustment". (TCER WorkingPaper Series, E-197.
- Nishimura, K. G and C. Shimizu (2003), "Distortion in Land Price Information—Mechanism in Sales Comparables and Appraisal Value Relation—," 東京大学日本経済国際共同研究センターディスカッションペーパー, No. 195.
- Shimizu, C. (2014), "Estimation of Hedonic Single-Family House Price Function Considering Neighborhood Effect Variables," *Sustainability*, 6, 2946-2960.
- Shimizu, C. and K. G. Nishimura (2006), "Biases in appraisal land price information: the case of Japan," *Journal of Property Investment & Finance*, 24(2), 150-175.
- Shimizu, C., K. Karato and Y. Asami (2010), "Estimation of Redevelopment Probability using

Panel Data-Asset Bubble Burst and Office Market in Tokyo-," *Journal of Property Investment & Finance*, 28(4), 285-300.

Shimizu, C., K. G. Nishimura and T. Watanabe (2016), "House Prices at Different Stages of Buying/Selling Process," *Regional Science and Urban Economics*, 59, 37-53.

国土庁土地局(1992),「土地情報の体系的整備体系について」国土庁.

国土庁土地局(1995),「諸外国の土地関連情報調査」国土庁.

国土庁土地局(2000),「土地市場における情報の欠如に伴う社会的損失に関する調査」国土庁.

清水千弘(1998),「不動産市場分析における統計的手法適用の最近の話題」不動産鑑定 1998.6月号」pp.77-91.

清水千弘・新井優太(2024),「マイクロデータを用いた計量経済学的接近法—ヘドニック法を用いた品質調整問題を題材として—」TEER Working Paper, J28.

東京都政策報道室(1999)『バブル期から現在に至る東京の土地市場の多角的分析』東京都.