

3. 相互補完的アプローチによる流通実態把握可能性

相互補完的アプローチ検討の基本的な考え方

■ 相互補完的アプローチの必要性

- 統計調査（新規統計も含む）とビッグデータ分析による流通実態データの把握手法は、次ページに示したような**特性の違いから単独で実施するには一長一短**がある。
- それぞれの単独手法の長所を活かし短所を補うためには、**両者を組み合わせた相互補完的アプローチが適当**と考えられる。
- 本章では、係る相互補完的アプローチ手法の技術的な選択肢を提示し、想定される手法の特性や課題を整理した。

■ 手法の実現可能性やデータ利用者ニーズへの対応可能性等を踏まえた検討

- 技術的選択肢として提示した相互補完的アプローチ手法のそれぞれについて、実現可能性やデータ利用者の多様なニーズを満たすデータを把握する手法としての適性を整理・評価した。
- 上記評価に基づき、**実現可能性が高く、かつ多様なニーズに答えられる手法として最も適していると考えられる相互補完的アプローチ手法を提案**した。
- 併せて、係る相互補完的アプローチ手法を用いても把握することの難しいデータ把握限界や今後の展望等についても検討した。

3. 相互補完的アプローチによる流通実態把握可能性 統計調査とビッグデータ分析の特性比較

	統計調査	ビッグデータ分析
データ収集可能頻度	調査票を使った調査(Webも含む)の場合は、記入者の負担を考えれば1～5年に1回程度が限界	月次、週次、日次など任意の頻度で収集可能
基本収集項目	量の把握を基本とするが、平均単価等から金額への換算も想定	- 現時点で想定されるビッグデータでは金額の把握が基本になると考えられる - 平均単価等を想定できれば量への換算も可能
収集データ粒度	項目ごとに集計された集計結果	取引伝票単位
データの偏在性	川下に向かうほど量は把握しにくくなる(歩留り、副産物、廃棄等の発生)	現時点では、POSデータ等のように川下が中心で、流通全体に対するカバー率は、川上に向かうほど小さくなると考えられる
サンプルの偏りや補正可能性	うまくサンプリングできれば、精度の高い補正と拡大推計の可能性は高い	- 個別企業が有するデータを分析するので偏りは大きい - 補正して拡大推計するためには全体構造を捉えた統計データが必要となる
データ収集に向けた問題点	調査協力者への作業負担が大きい	メリットやリスク等の観点から、データ保有者にとって提供のハードルが高い

想定される相互補完的アプローチ手法の技術的選択肢

	概要	手法の利点／展望	課題／論点
手法①	川上側を統計調査で捉え、川下側をビッグデータ分析でとらえる	- 流通実態を把握するためのデータの現時点での偏在性(統計調査データは川上ほど充実し、ビッグデータは川下ほど充実している)に即している - 現存するデータを最大限活かせる可能性が高い	- 川中のところで具体的にどのように両者を繋ぐのかは、その都度検討を要する - 川下側のビッグデータの偏りを補正するための統計データは別途必要となる
手法②	大きな流通構造全体を1～5年毎の統計調査で捉え、局所的な細かい変動を高頻度なビッグデータを分析することで捉える	- 調べたい局所のビッグデータが存在する(入手できる)場合は極めて有効 - 局所のビッグデータが存在する場合には、例えばコロナ禍における緊急事態宣言の影響等まで分析が可能	- どこに焦点を絞って局所的に分析するのかについては、その都度目的に応じて仮説を立てることが必要 - 調べたい局所のビッグデータが存在しない(入手できない)場合は不可
手法③	初期の段階では統計調査をメインとするが、徐々にビッグデータ分析の役割を拡大して将来的にはあらゆる段階の取引情報を電子化して提出することを義務付けしてビッグデータ分析をメインとする	早期にビッグデータを効果的に収集できる体制を整えられれば、あらゆる利用者ニーズに対して柔軟に対応できる	- 別途法制度化が必要 - 段階的に進めるとしても少なくとも20～30年にかかる可能性大 - インボイス制度の導入が追い風となる可能性も考えられる

想定される相互補完的アプローチ手法の技術的選択肢

	概要	手法の利点／展望	課題／論点
手法④	川上、川中、川下の「点」のデータを毎年統計調査し、それぞれの年変動の相関関係から「重み」係数を算出し、「線」のデータ自体は収集せずにこの「重み」を影響の強さとして代用する	「重み」に着目することで、各点の影響力の大きさが数値化できる - この「重み」が把握できれば「線」のデータ自体がなくても、例えばコロナや災害等でどこかに影響が出た場合の波及状況が概ね把握できる - 活用目的は限定されるが、「点」データを把握する既存統計だけでほぼ対応できる 「点」データがないものについてもニューラルネットワーク分析の中間層のような形で推計可能な場合もある	シンプルな流通構造であれば、現実に即して解釈できる可能性が高いが、流通構造が複雑であったり、似たような変動特性をもつ「点」が多数存在するような場合には、現実の流通構造と乖離する可能性もある。
手法⑤	②と④のハイブリッド ④で算出した重みの大きいところに焦点を絞って②のように局所的なビッグデータ分析を行う	手法②においてどこに焦点を絞って局所的に分析するかを検討をつけるのが課題となっていたが、そこに手法④を適用することで一定の示唆が得られると期待される。	焦点を絞りたい部分のビッグデータが必ずしも存在する(入手できる)とは限らない

想定される相互補完的アプローチ手法の技術的選択肢

	概要	手法の利点／展望	課題／論点
手法⑥	- 平時は統計調査で対応する（もしくはビッグデータを使う場合でも最新は必要としない・詳細項目は必要としない） - コロナ禍など緊急事態と認定された際には、特例としてビッグデータ保有元に最新時点・詳細項目の開示を求める	- 平時は統計調査だけでも対応可能 - 緊急時のみのビッグデータ開示であれば、保有者の開示抵抗感も低減できる可能性あり	- 緊急時の認定方法や開示義務等について法制度化も含めた議論が必要 - 緊急時のプロセスを十分に効率よく実行できるかが懸念 - 緊急時に急にビッグデータを用意しても正しい使用方法のナレッジが不足している可能性あり
手法⑦	- 政府統計を、1～数年に一回作成する - 週次や月次でビッグデータを用いた速報資料を作成する - 政府統計を作成するタイミングでビッグデータとも整合をとって統合版資料を作成する	- 週次や月次のビッグデータ速報資料については簡易版でも問題なし - 速報とすることでデータの正確性に関する最終的な保証を提供元に求めない - データ公表の壁がある場合、速報資料作成まで、ビッグデータ保有企業に委託する形式も考えられる	どのようなデータを速報資料として活用すべきか、過不足なく対応するためには多くの議論や交渉が必要
手法⑧	政府統計としてビッグデータを一次統計として活用する新しい流通実態加工統計（二次統計）※を作る	- ローデータとしてのビッグデータそのものは完全非公開とし、集計後の一次統計および加工した二次統計のみ公表とすればビッグデータ保有者の開示抵抗感を低減できる可能性あり - データ公表の壁がある場合、一次統計作成まで、ビッグデータ保有企業に委託する形式も考えられる	- 統計法にしっかりと位置付ける必要がある - その他の一次統計として政府統計調査を用いる場合、結果としてビッグデータの速報性などを喪失する可能性がある

※政府統計には、国民経済計算（GDP）や消費者物価指数など、他に存在する一次統計を複数集めて加工することで作成される加工統計もある

3. 相互補完的アプローチによる流通実態把握可能性 利用者ニーズに対するデータ把握手法の評価

評価軸		手法①		手法②	
		評価	評価のポイント	評価	評価のポイント
実現可能性	技術的難易度低い	△	・ 川中でのデータの繋ぎ方が難しい	○	・ 適切なビッグデータを入手できるかは不透明
	制度的難易度低い	◎	・ 基本的には現行制度の中で対応可能	◎	・ 基本的には現行制度の中で対応可能
データ特性に関する利用者ニーズへの対応可能性	統計学的信頼性	○	・ 川上部分に比べて川下部分の統計学的信頼性は低下する	○	・ 全体構造の統計学的信頼性は確保されるが、局所的なビッグデータ分析に高い統計学的信頼性は期待できない
	網羅性	△	・ 部分的にデータが欠落する可能性が高い	○	・ 全体構造部分の網羅性は期待できる
	全体を概略把握できる	△	・ 欠落したデータ部分の把握が難しい	◎	・ 全体構造は統計調査でしっかり把握できる
	局所変動を詳細把握できる	○	・ ビッグデータが使えれば、川下部分の局所変動については詳細把握可能	△	・ どこに焦点を絞るかの判断が難しい
	即時性	△	・ 川下部分のビッグデータ全てが日次または週次でなければならない	○	・ 焦点を絞った局所のビッグデータが日次または週次であれば可能
	地域性	△	・ 川下部分のビッグデータ全てに地域(地域間)が把握できる項目がなければならない	△	・ 取引情報等のビッグデータから地域を判断するのは困難
	今すぐ利用可能	○	・ 川下にはいくつか利用可能性のあるデータの存在が期待できる	○	・ 焦点を絞った部分にビッグデータがあれば可能
データ提供者の負担		○	・ 適正な対価があれば提供の可能性は高まる	○	・ 適正な対価があれば提供の可能性は高まる
コスト対効果		○	・ ビッグデータの入手コスト次第となる	○	・ ビッグデータの入手コスト次第となる
総合評価		△～○		○	

3. 相互補完的アプローチによる流通実態把握可能性 利用者ニーズに対するデータ把握手法の評価

評価軸		手法③		手法④	
		評価	評価のポイント	評価	評価のポイント
実現可能性	技術的難易度低い	○	・ 徐々に手法を改善していく余地がある	△	・ 複雑な現実の流通構造を反映したモデルを構築するのは難しい
	制度的難易度低い	△	・ 別途法制度化が必要で、提供義務化のハードルは高い	◎	・ 基本的には現行制度の中で対応可能
データ特性に関する利用者ニーズへの対応可能性	統計学的信頼性	○	・ ビッグデータ部分についても統計学的信頼性を確保しながら進める余地がある	△	・ モデルによる推計なので統計学的信頼性は想定外
	網羅性	○	・ 電子取引情報の提供が義務化されれば網羅性は高まる	△	・ 網羅的なモデルの構築は困難
	全体を概略把握できる	○	・ 電子取引情報の提供が義務化されれば全体を把握しやすい	○	・ 不足しているデータ等も含めて大局的な概略を把握できる可能性がある
	局所変動を詳細把握できる	○	・ 電子取引情報の提供が義務化されれば局所部分についても詳細分析が可能	△	・ 局所的な変動をモデルで再現するのは困難
	即時性	○	・ 日次・週次の電子取引情報が提供が義務化されれば即時性は確保できる	△	・ 即時性を持たせるには、モデルを構成するためのデータに即時性が必要
	地域性	○	・ 商流としての地域内・地域間変動は分析可能	△	・ 地域性を持たせるには、モデルを構成するためのデータに地域性が必要
	今すぐ利用可能	△	・ 初期段階では統計調査のみ利用可能	△	・ 適切なモデルの検討から始めなければならない
データ提供者の負担		△	・ 提供義務化の負担は大きい	○	・ 負担は既存統計調査と同等
コスト対効果		◎	・ 義務化されればデータ入手費用はほとんど想定されない	○	・ モデル構築の費用は不透明だが、うまくモデルが構築できれば効果が期待できる
総合評価		○		△	

3. 相互補完的アプローチによる流通実態把握可能性 利用者ニーズに対するデータ把握手法の評価

評価軸		手法⑤		手法⑥	
		評価	評価のポイント	評価	評価のポイント
実現可能性	技術的難易度低い	△	・ 複雑な現実の流通構造を反映したモデルを構築するのは難しい	○	・ 平時から緊急時にビッグデータを活用できるノウハウの蓄積が必要
	制度的難易度低い	◎	・ 基本的には現行制度の中で対応可能	△	・ 法制度化を含めた議論が必要
データ特性に関する利用者ニーズへの対応可能性	統計学的信頼性	○	・ 全体構造の統計学的信頼性は確保されるが、局所的なビッグデータ分析に高い統計学的信頼性は期待できない	○	・ 平時の統計調査は統計学的信頼性は確保されるが、緊急時の局所的ビッグデータ分析に高い統計学的信頼性は期待できない
	網羅性	○	・ 全体構造部分の網羅性は期待できる	○	・ 平時の網羅性は期待できる
	全体を概略把握できる	◎	・ 全体構造は統計調査でしっかり把握できる	○	・ 平時の統計調査である程度把握できる
	局所変動を詳細把握できる	○	・ 「重み」から局所変動個所を特定できれば有効	◎	・ 局所のビッグデータから詳細分析が可能
	即時性	○	・ 焦点を絞った局所のビッグデータが日次または週次であれば可能	○	・ 提供されたビッグデータの範囲では可能
	地域性	△	・ 取引情報等のビッグデータから地域を判断するのは困難	○	・ 提供されたビッグデータの範囲では可能
	今すぐ利用可能	△	・ 適切なモデルの存在が前提となる	○	・ 緊急時にビッグデータ入手ができれば可能
データ提供者の負担		○	・ 適正な対価があれば提供の可能性は高まる	○	・ 緊急時のみのビッグデータ提供は、提供者の負担軽減効果もある程度期待できる
コスト対効果		○	・ ビッグデータの入手コスト次第となる	○	・ 緊急時のみのビッグデータ入手なのでコスト低減効果はある程度期待できる
総合評価		○		○	

3. 相互補完的アプローチによる流通実態把握可能性 利用者ニーズに対するデータ把握手法の評価

評価軸		手法⑦		手法⑧	
		評価	評価のポイント	評価	評価のポイント
実現可能性	技術的難易度低い	◎	・ ビッグデータによる速報資料の作成や統合資料の作成について大きな問題は想定されない	○	・ 二次統計、一次統計、ビッグデータ間で連携した情報項目の設計が必要
	制度的難易度低い	○	・ ビッグデータの提供に関しては一部制度的な対応が必要になる可能性もある	△	・ ビッグデータ一次統計として活用するには、統計法にしっかりと日々日づける必要がある
データ特性に関する利用者ニーズへの対応可能性	統計学的信頼性	◎	・ 政府統計と統合するので、相応の統計学的信頼性は確保される	△	・ 一次統計の元になるビッグデータの偏り次第では信頼性の確保が困難
	網羅性	○	・ ある程度の網羅性は期待できる	○	・ ある程度の網羅性は期待できる
	全体を概略把握できる	○	・ 全体の概要把握はある程度期待できる	○	・ 全体の概要把握はある程度期待できる
	局所変動を詳細把握できる	○	・ ビッグデータのカバー範囲が、相応であれば週次・月次の速報分析により可能	△	・ ビッグデータの局所性は、加工段階で失われる可能性がある
	即時性	◎	・ 週次・月次の速報が有効	△	・ ビッグデータの即時性は、加工段階で失われる可能性がある
	地域性	○	・ 提供されたビッグデータの範囲では可能	○	・ 提供されたビッグデータの範囲では可能
	今すぐ利用可能	△	・ 必要なビッグデータの確保に向けた準備が必要	△	・ 必要なビッグデータの確保に向けた準備が必要
データ提供者の負担		○	・ ビッグデータの速報資料の作成まで想定すれば負担は増える	△	・ 提供義務化の負担は大きい
コスト対効果		○	・ ビッグデータの入手コスト次第となる	○	・ 義務化されればデータ入手費用は低減
総合評価		○～◎		△～○	

3. 相互補完的アプローチによる流通実態把握可能性 利用者ニーズに対するデータ把握手法の評価のまとめ

評価軸		手法①	手法②	手法③	手法④	手法⑤	手法⑥	手法⑦	手法⑧
実現可能性	技術的難易度低い	△	○	○	△	△	○	◎	○
	制度的難易度低い	◎	◎	△	◎	◎	△	○	△
データ特性に関する利用者ニーズへの対応可能性	統計学的信頼性	○	○	○	△	○	○	◎	△
	網羅性	△	○	○	△	○	○	○	○
	全体を概略把握できる	△	◎	○	○	◎	○	○	○
	局所変動を詳細把握できる	○	△	○	△	○	◎	○	△
	即時性	△	○	○	△	○	○	◎	△
	地域性	△	△	○	△	△	○	○	○
	今すぐ利用可能	○	○	△	△	△	○	△	△
データ提供者の負担		○	○	△	○	○	○	○	△
コスト対効果		○	○	◎	○	○	○	○	○
総合評価		△～○	○	○	△	○	○	○～◎	△～○

3. 相互補完的アプローチによる流通実態把握可能性 最適な相互補完的なアプローチの検討

■ 実現可能性が高く、最も多様なニーズに対応できるアプローチの検討

- これまで検討してきたように、想定される技術的選択肢としての各手法には一長一短があるが、今回設定した評価スキームに基づけば、最も高く評価されたのは以下の「手法⑦」である。

手法⑦の概要

- ・ 政府統計を、1～数年に一回作成する
- ・ 週次や月次でビッグデータを用いた速報資料を作成する
- ・ 政府統計を作成するタイミングでビッグデータとも整合をとって統合版資料を作成する

- ただし、どのようなデータを速報資料として活用すべきか、過不足なく対応するためには多くの議論や交渉が必要となり、手法の実現は中長期的な「目指すべき姿」としてとらえるべきものである。
- 今回検討してきた、別の技術的選択肢についても、各々に優れた点は存在し、データの使用目的が限定された場合には、別のアプローチ手法が最適となる可能性もある。
- また、これらの別のアプローチ手法は、目指すべき姿に到達する過程の、過渡期のアプローチとして適用可能性を検討する価値もある。
- **本検討の結論としては、上記手法を中長期的に目指すべき最適な相互補完的アプローチとして提案する。**

3. 相互補完的アプローチによる流通実態把握可能性 今後の展開に向けて

■ 流通実態に関するデータ把握の限界

- 流通実態をデータで把握しようとする場合、統計学的信頼性、網羅性、全体把握性、局所把握性、即時性、地域性等のあらゆる利用者ニーズに対応するデータを、限られた時間と予算の範囲で準備することには無理がある。
- 本検討においては、その中でも最も有用と考えられる手法として中長期的に目指すべきアプローチ手法を提案した。
- このような相互補完的アプローチが実現したとしても、依然として把握しきれないデータが残ってしまうことは否めない。
- 5品目の個別検討においては、統計調査におけるデータ把握の限界は、細かく見ていけば品目ごとに状況が異なることが判明しており、このような品目毎の状況に効率的に対応しつつ対象品目を拡大し、さらにビッグデータ分析等の有効なデータサイエンス的手法を活用しながらデータ把握の限界に挑み続けていくことが重要である。
- なお、データを把握するにあたり、ビッグデータ等の民間データを政府統計に活用することは、将来的には非常に有望ではあり、後述するように特に流通の寡占化・ボトルネック化が進んでいる領域においては特定少数の企業データと関連する取引先データを把握するだけで流通の全体像が把握しやすいメリットもある。ただし、政府統計として活用するためには、特定少数の企業データに頼るのはリスクがあることは留意しておく必要がある。

3. 相互補完的アプローチによる流通実態把握可能性 今後の展開に向けて

■ 今後の展開に向けた可能性の検討

【対象とする品目の拡大について】

- 今回詳細に検討したのは5品目であるが、本調査の結果としてこの5品目だけでもかなり流通構造が異なることが判明したことから、今後は品目の拡大に向けた具体的な検討が必要である。
- 今回の検討で、品目によっては寡占化・ボトルネック化の顕著なものなど比較的調査しやすい品目と、そうでない品目があること等が分かった。流通の寡占化・ボトルネック化が進んでいる領域においては、特定少数の企業データおよび関連する取引先データを把握するだけで流通の全体像が把握しやすいメリットがあるため、これらのメリットを活かして、まずは調査しやすいところに絞ってより詳細に掘り下げることにも有益である。
- 上記のような深掘り調査を複数品目で積み重ねていくことで、流通における寡占化・ボトルネック化の状況がある程度パターン化できる可能性がある。そうすれば、パターン毎にさらに最適な調査手法が見つかることも期待できる。
- また、品目を拡大していくにあたっては、「国民生活への影響が特に強いもの」「国内自給率が低く世界情勢に影響を受けやすいもの」など、ある程度優先順位をつけていくことも重要である。

【対象とする流通範囲の拡大について】

- 今回は「原材料」として形態を失わないところまでの流通を対象（加工品・飲食は原材料仕入れまで）としたが、川上から川下までのマテリアルフローを過不足なく把握する（最上流の物量合計と最下流の物量合計が一致する）ためには、加工段階の原材料構成比等も含めた検討が必要である。
- 川上は統計調査が川下はビッグデータが比較的充実しつつある中で、川中の製造加工部分が最も分かりにくい状況となっていることから、食品製造業や加工業は今後さらに検討を進める必要がある。
- 特に、災害対応等を想定した場合には、パン・菓子、総菜等の加工品の流通が極めて重要となる。
- また、災害等を想定した場合には、原材料・加工品だけでなく、包装材や生産機械等についても検討を進める必要がある。

今後の展開に向けて

■ 今後の展開に向けた可能性の検討(続き)

【対象とするデータ把握手法の拡大について】

- データ把握手法として今回主に検討したのは、統計調査手法、ビッグデータ分析を中心とするデータサイエンス手法、および両者を組み合わせた相互補完的アプローチ手法であるが、調査を通じて**ヒアリングによる流通構造の把握や不足データの収集・補完等が非常に有効で、今後もデータ把握手法の一翼を担うべき手法**であることも確認された。
- 今後は、統計調査、ビッグデータ分析、ヒアリングの三手法を有機的に組み合わせて、流通実態の把握に努めていくことが極めて有効であろうと考えられる。品目別に流通に詳しい専門学識者も多数存在することから、これらの研究者と日常的に情報交換していくことも重要である。
- また、データサイエンス手法として今回主に検討したのは、ビッグデータ分析とニューラルネットワーク分析であるが、**流通ネットワーク自体を対象とした分析としては、グラフ理論や複雑ネットワーク解析といった手法も有効**である可能性があり、今後の検討が望まれる。
- さらに、ビッグデータ分析のような比較的新しい手法だけでなく、**伝統的な産業連関分析等の手法も、不足部分の推計や地域間の推計等に活用できる**可能性も考えられるので、併せて検討していくことが望まれる。
- 産業連関分析を地域間で行うには、かつて政府統計として作成されていた地域間産業連関表のようなものが必要であり、調査頻度は5年に1回程度でも十分なので再整備されることが望まれる。
- 品目別の地域間産業連関表のようなものが整備されれば、特に物流問題等の解決には極めて有益である。

3. 相互補完的アプローチによる流通実態把握可能性 今後の展開に向けて

■ 今後の展開に向けた可能性の検討(続き)

【データの標準化について】

- 今後、食品流通全般においてビッグデータを有効に活用していくためには、現在のような加工品中心の製品コードではなく、生鮮食品、日配品、総菜等にも簡単に対応できるような**標準コード体系を確立し、普及させていくことが重要**である。

【データ提供の根拠とインセンティブについて】

- 速報性を重視するには民間のビッグデータ利用が不可欠になると考えられるが、現時点では法的な提供根拠がないため、個別に提供依頼をして承諾していただくことが前提となる。
- そのため、**データを提供する事業者にとの様なメリット・デメリットがあるかについては、さらに整理していく必要がある**。また、データを扱っている企業のデータは、自分のデータではないこともあり、事業者のリスクや負担もある。今後に向けては、これらの問題や課題をさらに明確にしていくことが重要である。
- ビッグデータの提供については、**将来的にどのような形であれば法的な提供を義務化すること等が可能になりうるのか**等の議論を、可能な限り早急を開始することが必要である。