図書館クラウディング・アウト効果と公共貸与権

全国図書館大会 11月22日 三重県立文化会館

発表者：川口康平　香港科技大商学院経済学部

共著者：金澤匡剛　東京大学大学院経済学研究科

まとめ

- 2017年の市区町村・月・書籍レベルの書店売上、図書館蔵書データを用いて、図書館蔵書の書店売上に対する**平均的因果効果**を推定した。
- ある市区町村で図書館蔵書が一冊増えると、書籍の売上は平均して毎月**0.12**冊ほど減る。
- その効果は人気の書籍ほど高く、上位1/6以外の書籍についてはそうした効果はほとんどない。
- 図書館蔵書全体の存在による書籍の逸失売上は**17.5%**ほど。
- ただし、図書館蔵書が書店売上を奪うとしても、そのことが直ちに公共貸与権の必要性を導くわけではない。

因果推論とは何か

XのYに対する「平均的」因果効果

- **定義：**<u>他の条件の平均を一定として</u>、Xを一単位変化させたときにYが何単位変化するか。
- 政策によってXを変化させるという介入を行うときにはこの「平均的」因果効果に関する見積もりを持っておくことが望ましい。
- ある種の理念型であって測定は容易ではない。

少人数学級の効果の例

- 例：小学校のクラスの人数を少なくすると子供の成績は良くなるか？
- 市内に小学校が20校ある。
- 小学校群A（統制群）：今まで通り一クラス40人×二クラス×10校で計800人。
- 小学校群B（治験群）：一クラス20人×四クラス×10校で計800人。
- この状態で一年間授業。年度終わりに統一テストを実施。
- 統計量＝小学校群Bのテストの点数の平均点－小学校群Aのテストの点数の平均点。

少人数学級の効果の例

- ある生徒のテストの点数＝その学校の教師の資質＋その生徒の資質＋**少人数学級の因果効果**。
- 小学校群Bのテストの点数の平均＝小学校群Bの教師の資質の平均＋小学校群Bの生徒の資質の平均＋少人数学級の**平均的因果効果**。
- 小学校群Aのテストの点数の平均＝小学校群Aの教師の資質の平均＋小学校群Aの生徒の資質の平均。
- 統計量＝（小学校群Bの教師の資質の平均－小学校群Aの教師の資質の平均）＋（小学校群Bの生徒の資質の平均－小学校群Aの生徒の資質の平均）＋**少人数学級の平均的因果効果**。
- 最初の２つの項が打ち消し合って0にならない限りこの統計量は少人数学級の平均的因果効果を表していない。

バイアスの例

- 小学校群Bの教師の資質の平均－小学校群Aの教師の資質の平均＞ 0
  - 小学校群Bは少人数学級に効果があると見せかけようとしてこの学年に学内でも特に優秀な先生を割り当てた。
  - 上方バイアス。
- 小学校群Bの教師の資質の平均－小学校群Aの教師の資質の平均＜0
  - 小学校群Aは自分の学校の成績が比較されて悪く見えるのを恐れてこの学年に学内でも特に優秀な先生を割り当てた。
  - 下方バイアス。

バイアスの例

- 小学校群Bの生徒の資質の平均－小学校群Aの生徒の資質の平均＞ 0
  - 少人数学級の効果を大きく見せるためにその市の教育委員会は「良い学校」と認知されている小学校で少人数学級を実施した。
  - 上方バイアス。
- 小学校群Bの生徒の資質の平均－小学校群Aの生徒の資質の平均＜0
  - その市の教育委員会は成績を改善したい学校に優先して少人数学級を導入した。
  - 下方バイアス。

統計的因果推論の信頼性

- 因果推論の信頼性は「他の条件の平均を一定として」という条件をどの程度保証できるかにかかっている。
- 「どの小学校に少人数学級を実験的に導入するかはランダムに決めたので統制群と治験群で生徒の資質は平均的に同じになっているはず」
- 「どの小学校に少人数学級を実験的に導入するかを決定する前に教員の割当計画を提出させ、導入決定後は割当計画の変更を認めなかったので統制群と治験群で教員の資質は平均的に同じになっているはず」
- **すべての懸念に対する完全な保証は不可能。**
- **「具体的な懸念→具体的な保証」をどれだけ積み重ねられるかが鍵。**

統計的因果推論の信頼性

- 喩えて言えば、**統計的因果推論の信頼性＝ポーカーの手の強さ**。
- 「具体的な懸念に対する具体的な保証」の組み合わせは役のようなもの。
- ストレートが揃っていても100%勝てるわけではないが、普通は勝負する。
- ワンペアだけでも100%負けるわけではないが、普通は勝負しない。
- ポーカーで勝負に出るか否かを決めるにあたって手の強さを考慮するように、政策を実施するか否かを決めるにあたっても統計的因果推論の信頼性を考慮する。

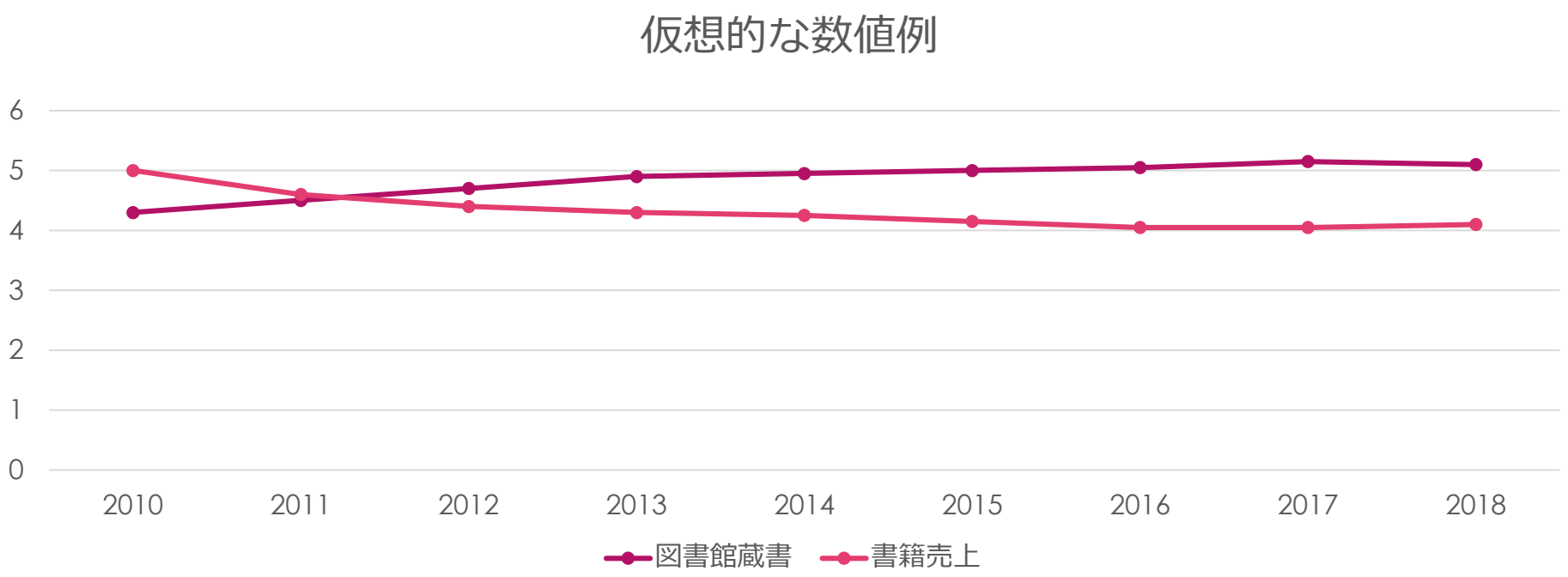
図書館蔵書のクラウディングアウト効果

既存の分析の問題点

図書館蔵書のクラウディングアウト効果

- 「**図書館蔵書(X)の書籍売上(Y)に対する平均的因果効果**」を測定したい。
- ある市区町村である書籍の蔵書数を一冊増やしたときにその市区町村の書店でのその書籍の売上は変化するか？変化するとしたら増加するか？減少するか？何単位変化するか？
- 既存の分析の信頼度はポーカーで言えば「ハイカード」程度。
- 私達の論文の分析の信頼度は「スリーカード」程度。

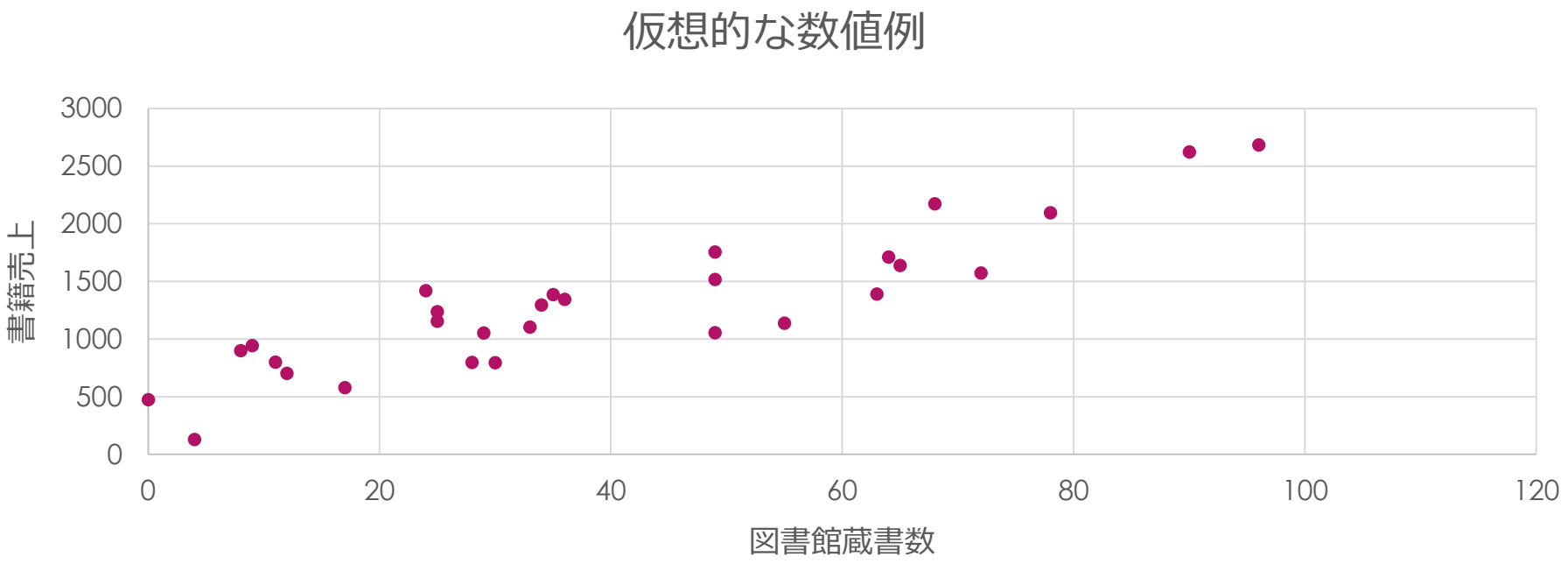
時系列比較


仮想的な数値例
6
5
4
3
2
1
0
2010
2011
2012
2013
2014
2015
2016
2017
2018
図書館蔵書
書籍売上

時系列比較の問題点

- 「他の条件の平均」は年度間で一定か？
- 「図書館蔵書はストックなので廃棄はあれど総蔵書数は維持されるか微増する」
- 「時間消費が書籍からネットに移行しているので書籍売上は減少していく」
- これらのトレンドの結果、図書館蔵書↑書籍売上↓のように見える。
- 時系列比較はこのような懸念に答えない。

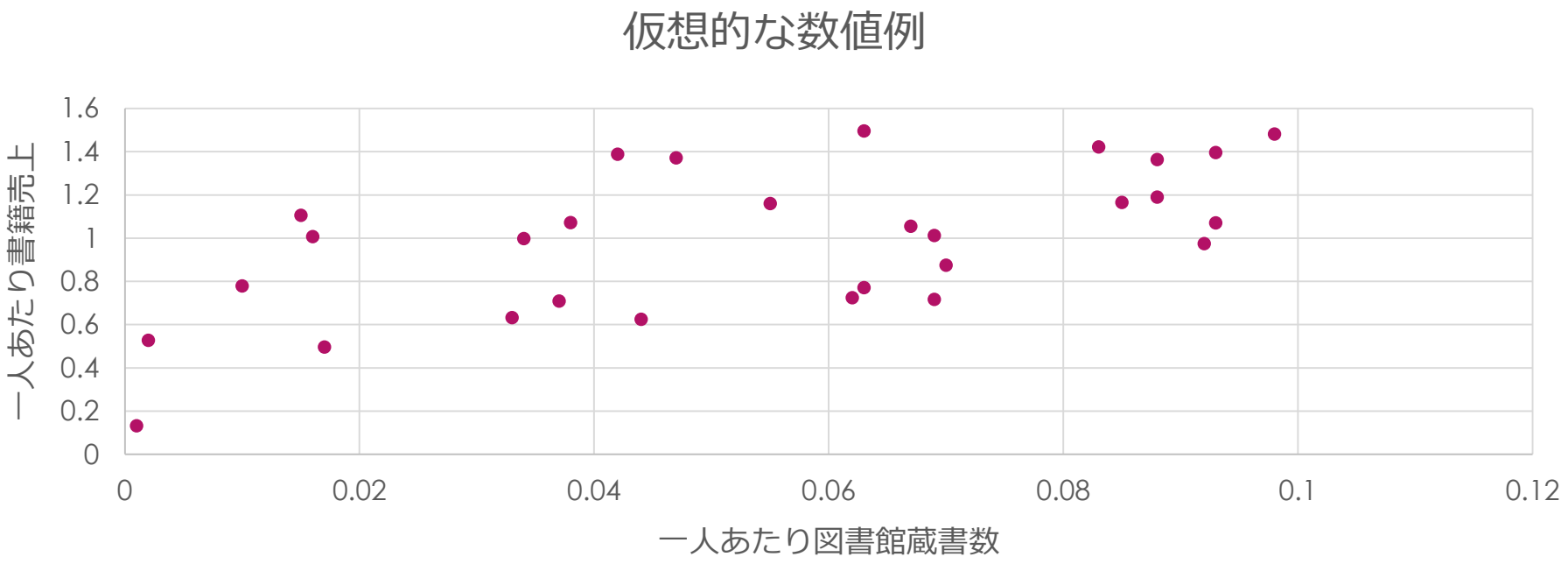
書籍間比較


仮想的な数値例
3000
2500
2000
1500
1000
500
0
書籍売上
0
20
40
60
80
100
120
図書館蔵書数

書籍間比較の問題点

- 「他の条件の平均」は書籍間で一定か？
- 書籍間でそもそもの人気度が違うのではないか？
- 「人気の書籍はいろいろな地域で平均的にたくさん売れる」
- 「人気の書籍はいろいろな地域の図書館で平均的に蔵書されやすい」
- これらの傾向の結果、図書館蔵書数↑書籍売上↑に見える。
- 書籍間比較はこのような懸念に答えない。

市区町村間比較


仮想的な数値例
一人あたり書籍売上
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.02
0.04
0.06
0.08
0.1
0.12
一人あたり図書館蔵書数

市区町村比較の問題点

- 「他の条件の平均」は市区町村間で一定か？
- 市区町村間でそもそもの読書傾向がちがうのではないか？
- 「読書好きの市区町村では一人あたり図書館蔵書数が多い」
- 「読書好きの市区町村では一人あたり書籍売上が多い」
- これらの傾向の結果、図書館蔵書↑書籍売上↑に見える。
- 市区町村比較はこのような懸念に答えない。

時系列分析、横断的分析の問題点

- これらの問題は時系列分析、横断的分析一般の問題。
- 統制変数（書籍の著者、市区町村の若年人口比率、等）を増やして他重回帰分析を行えば「それらの統制変数の平均の違い」は一定にできる。
- しかし、統制変数以外の条件の平均の違いは統制できない。
- サンプルサイズを増やすと平均値がより精確に測定できるようになるが、「他の条件の平均」がそれによってなくなるわけではないので、サンプルサイズを増やしても統計的推論の信頼度は向上しない。

図書館蔵書のクラウディングアウト効果

縦断的分析

縦断的分析

- 信頼できる統計的因果推論の**最低ライン**。
- 「時点間のその他条件の平均の違い」「書籍間のその他条件の平均の違い」「市区町村間のその他条件の平均の違い」に対する懸念**すべて**に答える。

縦断的分析のアイディア

- ある書籍のある市区町村でのある月の売上＝
  - **その書籍のその市区町村でのその月の図書館蔵書の影響**
  - ＋<u>その書籍のその市区町村での人気</u>＋<u>その書籍のその月の人気</u>
- …だったとする。
- **太線部**がみたい平均的因果効果。<u>下線部</u>が統制したいその他条件。
- その書籍のその市区町村での人気。
  - 書籍の著者、市区町村の人口構成などさまざまな要因が影響する。
- その書籍のその月の人気。
  - 書籍の宣伝、口コミなどさまざまな要因が影響する。

縦断的分析のアイディア

- ある書籍のある市区町村での５月の売上＝**その書籍のその市区町村で５月の図書館蔵書の影響**＋その書籍のその市区町村での人気＋その書籍の５月の人気
- ある書籍のある市区町村での４月の売上＝**その書籍のその市区町村で４月の図書館蔵書の影響**＋その書籍のその市区町村での人気＋その書籍の４月の人気
- 上の式から下の式を引くと…
- ある書籍のある市区町村での５月の売上－ある書籍のある市区町村での４月の売上
  - ＝（**その書籍のその市区町村で５月の図書館蔵書の影響－その書籍のその市区町村で４月の図書館蔵書の影響**）＋（その書籍のその市区町村での人気－その書籍のその市区町村での人気）＋（その書籍の５月の人気－その書籍の４月の人気）
  - ＝（**その書籍のその市区町村で５月の図書館蔵書の影響－その書籍のその市区町村で４月の図書館蔵書の影響**）＋（その書籍の５月の人気－その書籍の４月の人気）

縦断的分析のアイディア

- ある書籍の市区町村Aでの５月の売上－ある書籍の市区町村Aでの４月の売上
  - ＝ （**その書籍の市区町村Aで５月の図書館蔵書の影響－その書籍の市区町村Aで４月の図書館蔵書の影響**）＋（その書籍の５月の人気－その書籍の４月の人気）
- ある書籍の５月の市区町村間平均売上－ある書籍の４月の市区町村間平均売上
  - ＝ （**その書籍の５月の図書館蔵書の影響の平均－その書籍の４月の図書館蔵書の影響の平均**）＋（その書籍の５月の人気－その書籍の４月の人気）
- 上の式から下の式を引くと…

縦断的分析のアイディア

- （ある書籍の市区町村Aでの５月の売上－ある書籍の市区町村Aでの４月の売上）－（ある書籍の５月の市区町村間平均売上－ある書籍の４月の市区町村間平均売上）
  - ＝ **（その書籍の市区町村Aで５月の図書館蔵書の影響－その書籍の市区町村Aで４月の図書館蔵書の影響）－（その書籍の５月の図書館蔵書の影響の平均－その書籍の４月の図書館蔵書の影響の平均）**＋（その書籍の５月の人気－その書籍の４月の人気）－（その書籍の５月の人気－その書籍の４月の人気）
  - ＝ **（その書籍の市区町村Aで５月の図書館蔵書の影響－その書籍の市区町村Aで４月の図書館蔵書の影響）－（その書籍の５月の図書館蔵書の影響の平均－その書籍の４月の図書館蔵書の影響の平均）**
- …と「その他条件の違い」に起因する項が消せる。

縦断的分析のアイディア

- 各書籍について…
- 同市区町村内の4月と5月の売上の変化はその市区町村における時間変化しない特性（人口特性など）には影響を受けない。
- したがって「同市区町村内の変化」に着目すれば、その市区町村の時間変化しないその他条件の影響は無視できる。
- 「市区町村A内の4月と5月の売上の変化」と「4月と5月の平均売上の変化」の差はその書籍の4月と5月の特性の変化（全国的な宣伝、口コミなど）には影響を受けない。
- したがって「市区町村A内の売上変化と市区町間平均売上の変化の差」に着目すれば、その書籍の4月と5月の特性に関するその他条件の影響は無視できる。

縦断的分析のアイディア

- 言い方を換えると…
- 「ある書籍に関してある月から次の月にかけて、ある市区町村でのその書籍の図書館蔵書数が、その月から次の月にかけてのその書籍の図書館蔵書数の全市区町村での平均的な変化よりも一単位多く変化したとき、その市区町村でのその書籍の書店売上は、その月から次の月にかけてその書籍の書店売上の全市区町村での平均的な変化よりも何単位多く変化するか？」
- …に着目すると、
  - ある書籍のある市区町村での人気に影響する様々な要因
  - ある書籍のある月における人気に影響する様々な要因
- の影響を**すべて**取り除いて平均的因果効果を測定することができる。

仮想的な数値例

- 2019年10月から11月にかけて各市区町村では平均して図書館蔵書数が0.5から1.5に増えた。同時期、書店売上は平均して5から8に増えた。
- 一方、三重県津市では図書館蔵書が0から2に増え、書店売上は4から6に増えた。
- 三重県津市では図書館蔵書が平均と比較して(2-0)-(1.5-0.5)=1増えたときに、書店売上は平均と比較して(6-4)-(8-5)=-1変化した。
- 図書館蔵書の平均以上の伸びによって、書店売上の伸びが平均以下に抑えられた。
- 図書館蔵書のクラウディングアウト効果あり。

縦断的分析に残る問題点

- ある書籍のある市区町村でのある月の売上＝
  - **その書籍のその市区町村でのその月の図書館蔵書の影響**
  - ＋その書籍のその市区町村での人気＋その書籍のその月の人気
- この仮定からはずれる懸念には答えられない
- 例：その市区町村ではその月にたまたまその書籍の著者がサイン会にやって来ることになっていて、そのことを予め知っていた図書館側が蔵書を増やしていたら？
- 「ある市区町村で、ある書籍について、ある月にだけ図書館蔵書と書店売上が平均以上に同時に増えたり減ったりする要因」が**広汎に**みられるならこの縦断的分析には問題がある。
- あるだろうか？

図書館蔵書のクラウディングアウト効果

分析結果

統計的平均因果効果

- 平均的には、ある書籍について、ある市区町村である月からその次の月にかけて、その書籍の図書館蔵書数がその月から次の月への平均的な変化と比べて1単位余分に増えた場合、その市区町村でその月からその次の月にかけて書店売上はその書籍の平均的な変化と比較して**0.12冊**減少する。
- (分析は4-7月のデータだが、1-2月の売上が)上位1/6に属する書籍に関しては**0.36冊**減少する。
- 2015、2016年度紀伊国屋ベストセラーTop50書籍の場合は**0.62冊**減少する。
- 一年を通じて、この10倍程度の売上が奪われる。

図書館蔵書による書店の逸失売上

- 以上の推定結果を踏まえて、仮に2017年4-7月に図書館蔵書が全く無かった場合、書店の売上はどの程度増えていたか？
- 実際の売上の**17.5%**ほど増えていた、という概算値が得られる。
- 図書館が書籍購入に使っている支出は書籍売上全体の5%程度なので、ある程度の逸失売上をもたらしているといえる。

残る懸念は？

- 縦断的分析では、その書籍・その市区町村・その月に図書館蔵書数にも書店売上にも影響するようなその他条件の影響（サイン会など）に関する懸念には答えられない。
- しかし、そうした影響のほとんどは、サイン会の例のように図書館蔵書↑書店売上↑になるものであると考えられる。
- その場合は、平均的因果効果の測定値には**上方バイアス**がかかる。
- これはつまり、本来は0.12冊よりも売上が減っているにも関わらず0.12冊程度しか減っていないように見えるように働く、ということ。
- それゆえ、今回の分析結果は図書館蔵書のクラウディング・アウト効果の**保守的な見積もり**である、と解釈している。

よく聞かれる質問

- 図書館蔵書の書店売上への影響ではなく、図書館貸出の書店売上への影響をみるべきでは？
  - 消費者は「その本を読まない、その本を買って読む、その本を借りて読む」のうちどれを選択するかを考える。
  - 図書館蔵書数はこの選択問題を決定づける変数。
  - 消費者行動を分析するときは、環境がどの選択を導くかを見る。
  - 「本を借りて読んだからかその本は買わなかった」というのは、選択肢のあいだの代替性から導かれる論理的な関係であって、因果効果ではない。
  - なので、書店売上を図書館貸出に回帰してもその分析に意味はない。

よく聞かれる質問

- 図書館蔵書が書店売上を減らすのであれば、やはり公共貸与権は必要か？
  - 図書館蔵書が書店売上を減らすことが**直ちに**公共貸与権の必要性を導くわけではない。
  - 特許や著作権による権利の保護の程度は、その保護によって発明や著作を行うインセンティブを生み出すというメリットと、その保護によってその発明や著作の利用の機会を奪うというデメリットを天秤にかけて決めるもの。
  - 今回見いだされた著者への損失はどの程度著作のインセンティブを減らしているのか？仮に図書館に著者の損害への補填を求めた場合図書館蔵書はどの程度減るか？その結果図書の利用機会はどの程度減るか？…といった議論なしには決定できない。
  - そうした大枠の議論を離れて、単純に著者への一律の補填を行うのであれば、保護月数×価格の12%程度の補填を行うのがよいと思われる。

まとめ

- 2017年の市区町村・月・書籍レベルの書店売上、図書館蔵書データを用いて、図書館蔵書の書店売上に対する**平均的因果効果**を推定した。
- ある市区町村で図書館蔵書が一冊増えると、書籍の売上は平均して毎月**0.12**冊ほど減る。
- その効果は人気の書籍ほど高く、上位1/6以外の書籍についてはそうした効果はほとんどない。
- 図書館蔵書全体の存在による書籍の逸失売上は**17.5%**ほど。
- ただし、図書館蔵書が書店売上を奪うとしても、そのことが直ちに公共貸与権の必要性を導くわけではない。