

インターネット調査の偏りと補正

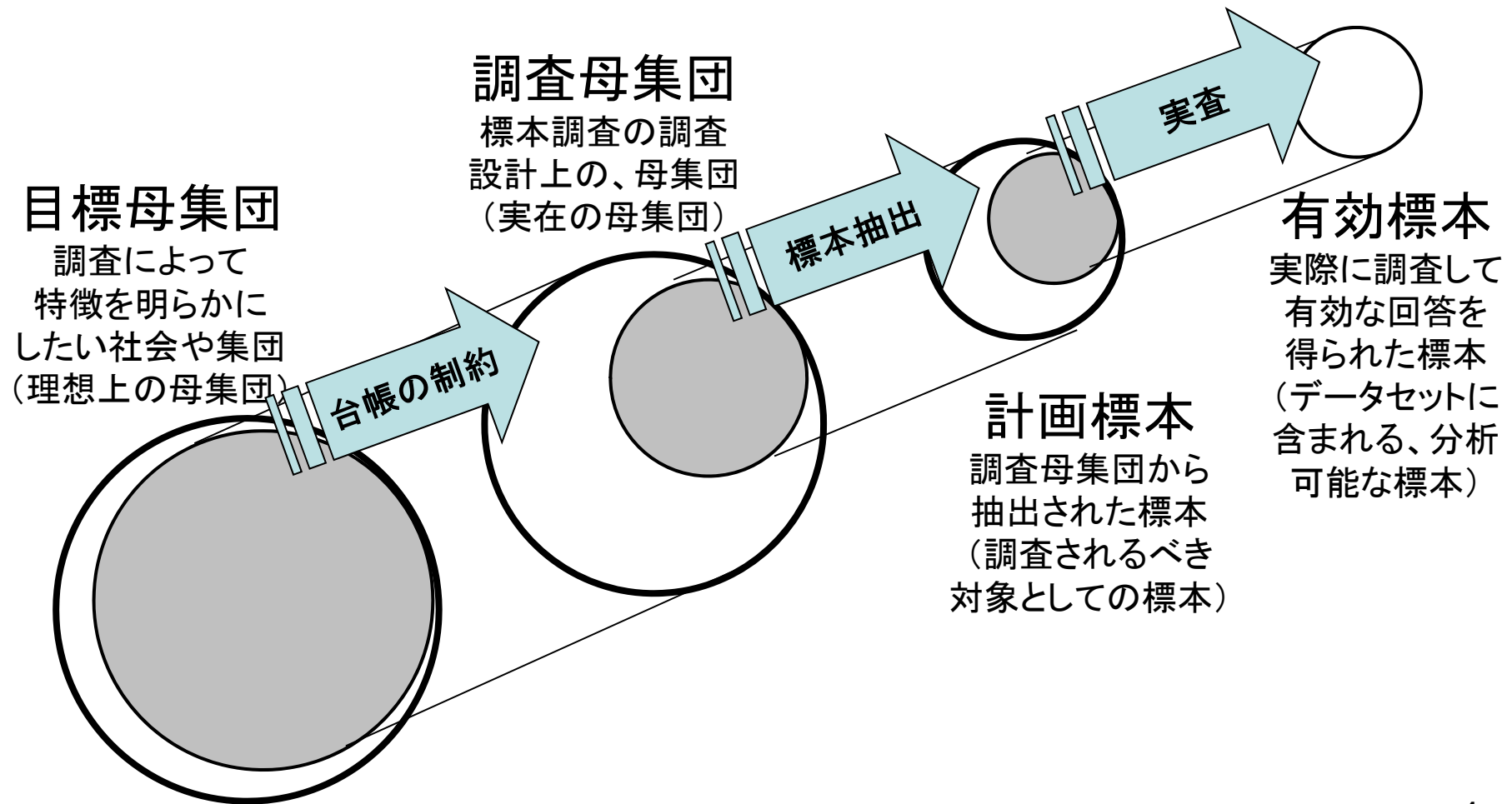
三輪 哲
(東北大学)

全体構成

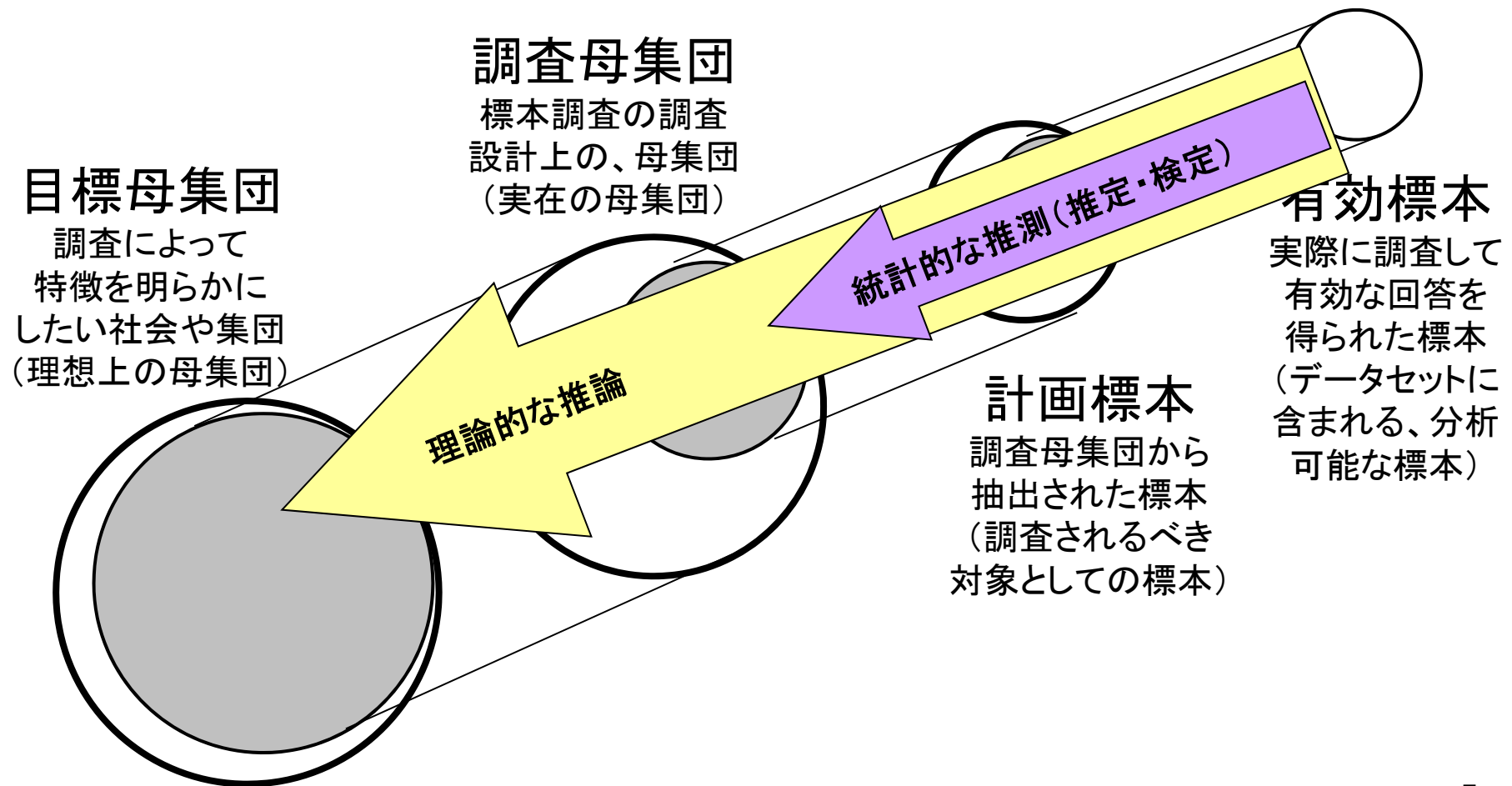
- 標本調査における誤差
- 比較調査設計と全体的な差異
- 答える人の違い
- 答え方による違い
- 補正の可能性

標本調査における誤差

対象者を選ぶ



対象者を選ぶ



標本調査データの正確さ

- 標本を調べるだけで、全体（母集団）のことはある程度わかる
 - 重要なのは、「どの程度正確か」
 - 誤差の評価が重要になる
 - 誤差にはいくつかの種類がある
- 「誰が答えるか」に関わる誤差の種類
 - カバレッジ誤差、標本誤差、非回答誤差

カバレッジ誤差

- カバレッジ誤差：
 - 母集団に含まれているのに、標本にどうしても選ばれない対象者が存在することに起因する誤差
 - 「標本に含まれる可能性がある人々」と「標本に含まれる可能性がない人々」との間に違いがある場合に生じる
 - 「どの集団が母集団か」を考えないと始まらない(！)

カバレッジ誤差

- 例:「日本の全住民」の内閣支持率を調べるときに...
 - 選挙人名簿から無作為に対象者を選ぶ
 - 全住民が台帳に載っているなら、カバレッジ誤差なし
 - 無作為な固定電話番号に電話して対象者を選ぶ
 - 固定電話を持たない人が標本に含まれることがないので、カバレッジ誤差が生じる可能性

標本誤差

- 標本誤差：母集団ではなく、その一部である標本だけを調べることに起因する誤差
- 標本に含まれる人を「サイコロ」で100人選ぶとき
...
 - 1回目に選んだ100人では内閣を支持する人が25%かもしれない
 - 「サイコロ」を振りなおして、2回目に選んだ100人では、内閣を支持する人が15%かもしれない
 - 「サイコロ」の振りなおしでどれぐらい結果が変わってくるか＝標本誤差
- 通常、統計的手法が評価する「誤差」は標本誤差のみを指す

標本誤差

- 標本誤差は統計的手法で評価できる
 - 標本の無作為性が仮定される
- 適切に評価できるとしても、やはり標本誤差は小さい方がよい
 - 標本誤差が大きいと、観察された差が「誤差の範囲」なのか「本当の差」なのかを判別できない(検出力が弱い)
 - 標本サイズ(標本に含まれる人の数)が大きい方がよい

非回答誤差

- 非回答誤差：
 - 標本として選ばれた人のうち、調査に回答しない人が存在することに起因する誤差
 - 「回答する人々」と「回答しない人々」との間に相違がある場合に生じる

非回答誤差

- 例：内閣支持率を調べるときに...
 - － 標本として選ばれた人がすべて回答した
→ 非回答誤差なし
 - － 回答しない人もいたが、「回答した人々」と「回答しない人々」との間に支持率は異なる
→ 非回答誤差なし
 - － 内閣を支持する人々が回答しやすい傾向がある
→ 非回答誤差あり

誤差と偏り

- 誤差
 - プラス方向とマイナス方向へと影響
 - 系統だった差は生じさせない
 - 推定値の精度が低下する
- 偏り
 - どちらか一方向へと影響
 - 系統だった差が生じてしまう
 - 推定値に体系的な誤りが生じる

具体的な留保事項

- インターネット調査のデータ
 - 全体をカバーできるか？
 - × 登録型モニター
 - × 公募型
 - 無作為な標本抽出ができるか？
 - 前提となるのは母集団のリストの存在
 - 非回答を評価できるか？
 - × 先着順

「やる気のある人」が回答するための偏り

比較調査設計と全体的な差異

比較調査の設計(萩原 2009より)

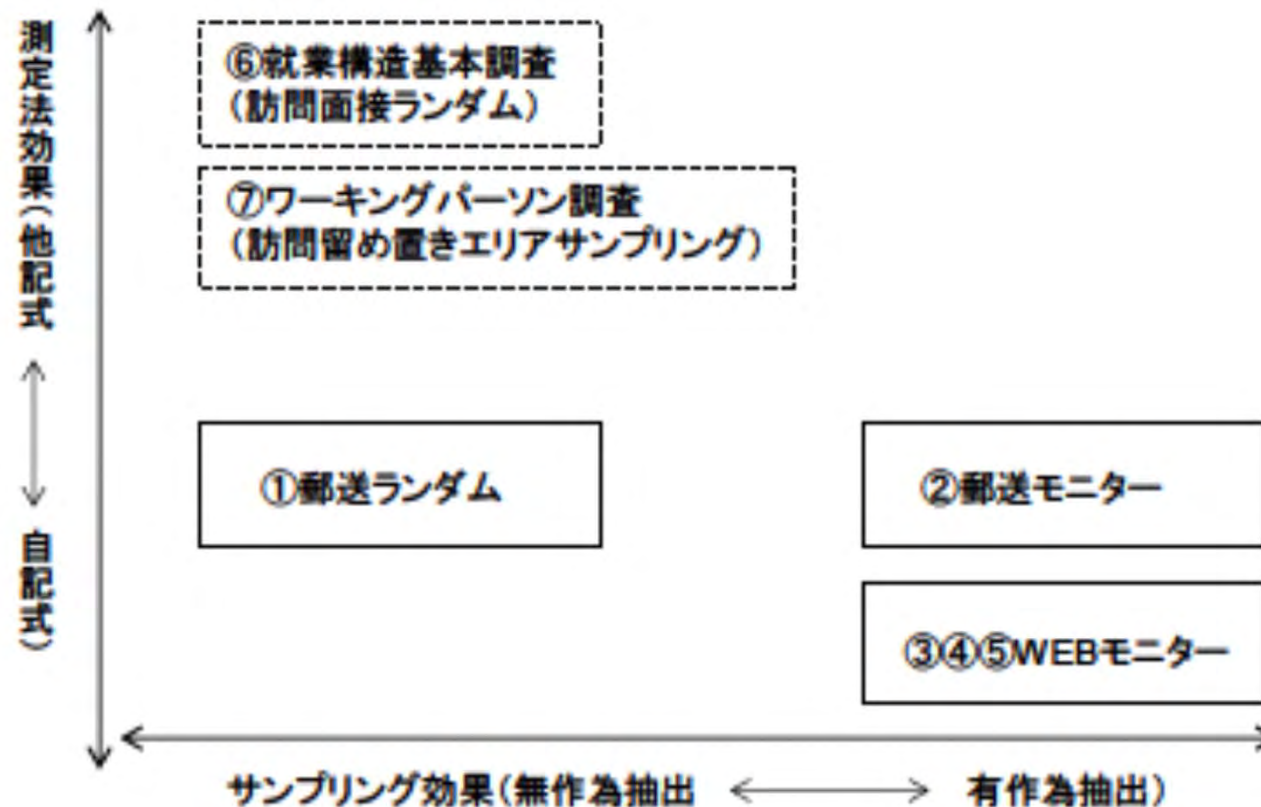


図 1 調査の基本設計

比較調査の設計(萩原 2009より)

表 3 調査設計詳細

調査名	配信数	回収数	回収率	回収締め切り方法	催促の有無	依頼主
①郵送ランダム	3500	753	21.5%	調査期間終了後、予定より多く回収したセルは無作為抽出	10/16に督促はがき発送	開示
②郵送モニター	1340	1108	82.7%		10/16に督促はがき発送	非開示
③WEBモニターA	2666	1201	45.0%		督促なし。回収数未達のセルのみ追加配信10/15に実施	
④WEBモニターB	2005	1472	73.4%	先着順受付	督促は1回(10/16) 配信方法は回収数をみながら 期間中分割配信	非開示
⑤WEBモニターC	29995	2088	7.0%	調査期間終了後、予定より多く回収したセルは無作為抽出	10/15に1回	

インターネット調査の偏り

- 答えている人が違うのでは？
 - 調査選択バイアス
 - どういう調査法にするかによって、回答する層が変わる
 - モニター効果
 - 同じ調査法でも、モニターかランダムサンプルかで、違う
- 答え方によって違いが出るのか？
 - モード効果
 - 同じ人でも、紙に書くのと、PCで入力とで、違う

MDSの項目の布置

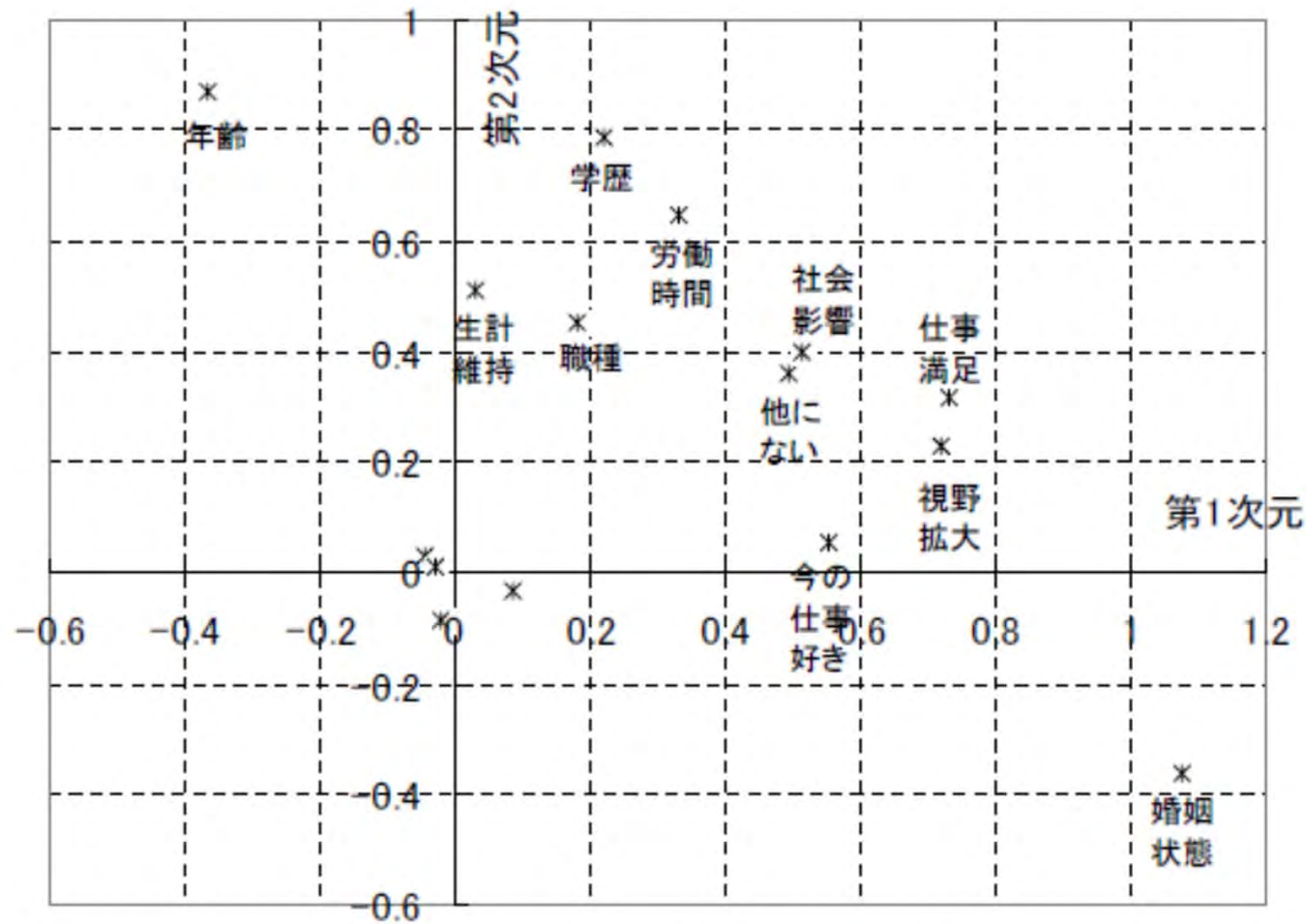


図6 個人差MDSの個人(項目)付置(第1次元×第2次元)

MDSの項目の布置

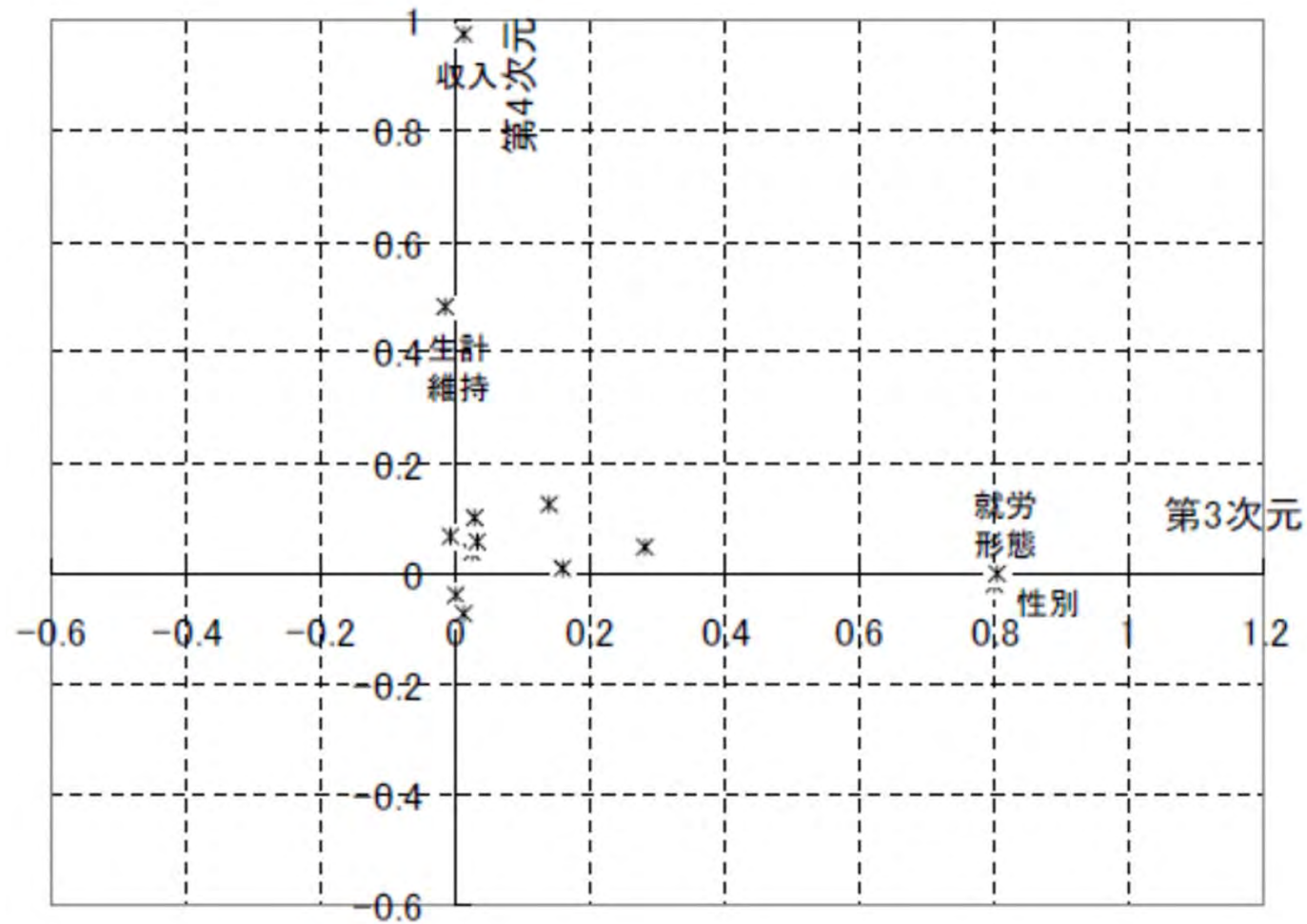


図7 個人差MDSの個人(項目)付置(第3次元×第4次元)

MDSの共通対象布置

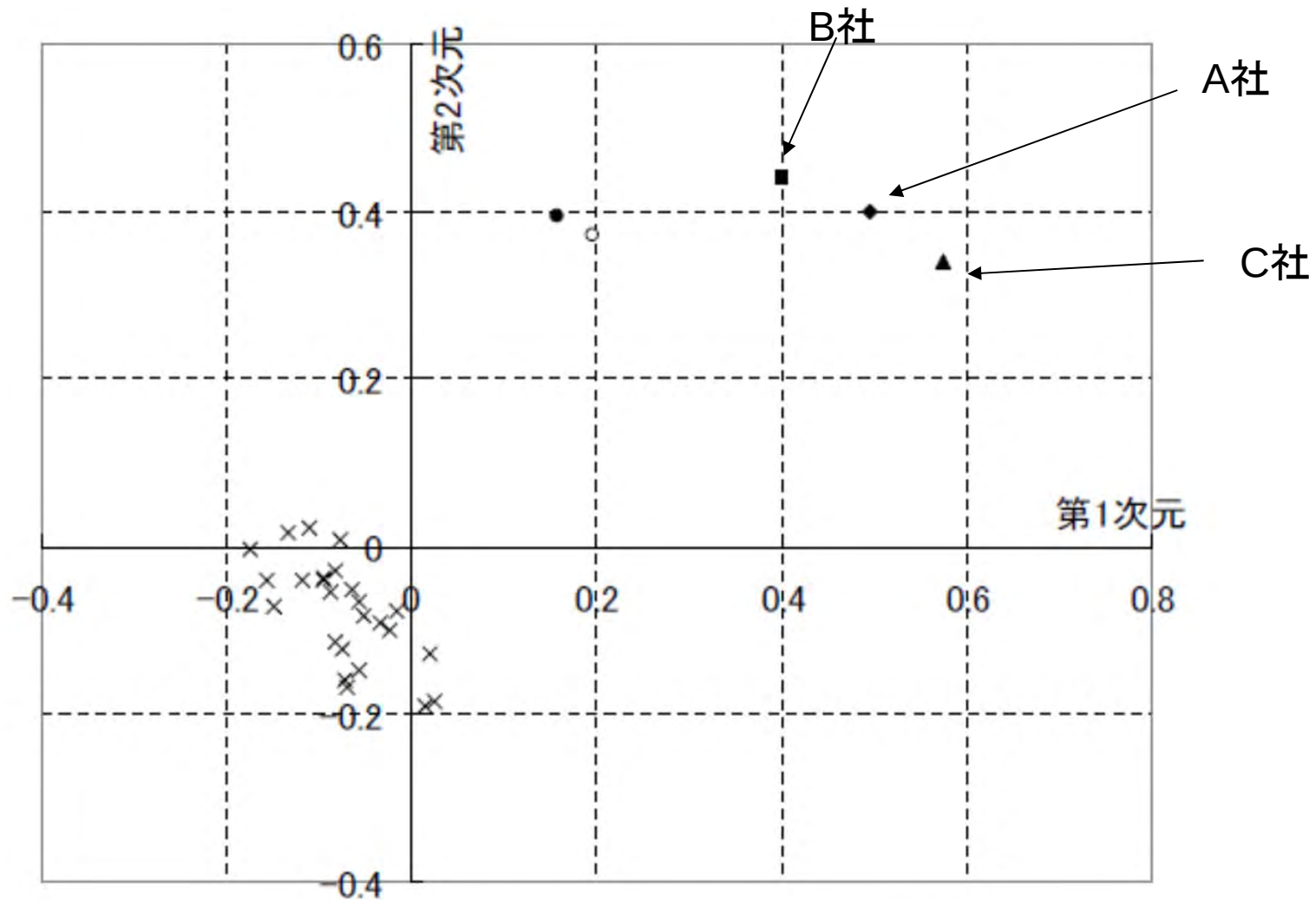


図1 個人差MDSの共通対象付置(第1次元×第2次元)

MDSの共通対象布置

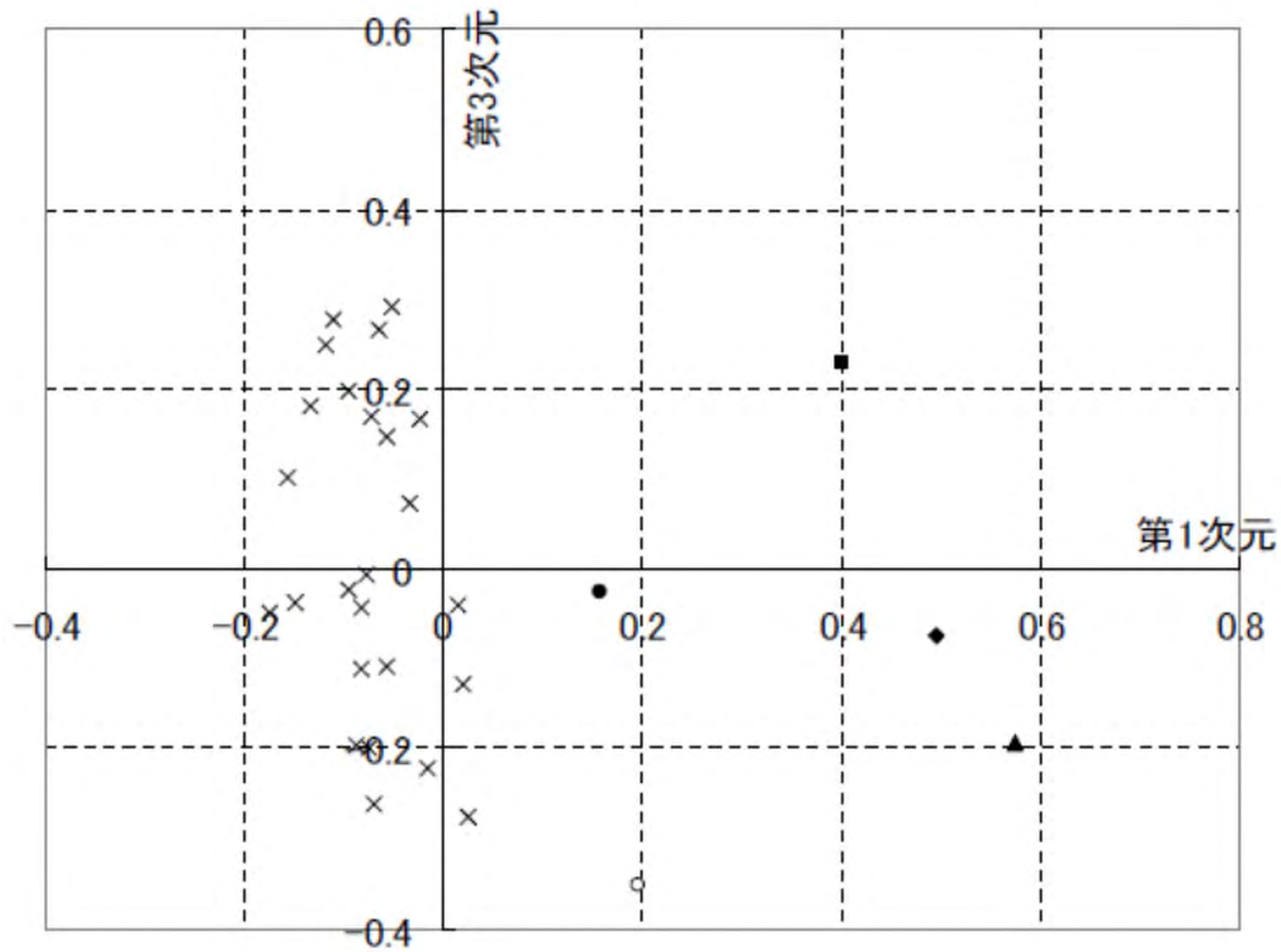


図2 個人差MDSの共通対象付置(第1次元×第3次元)

MDSの共通対象布置

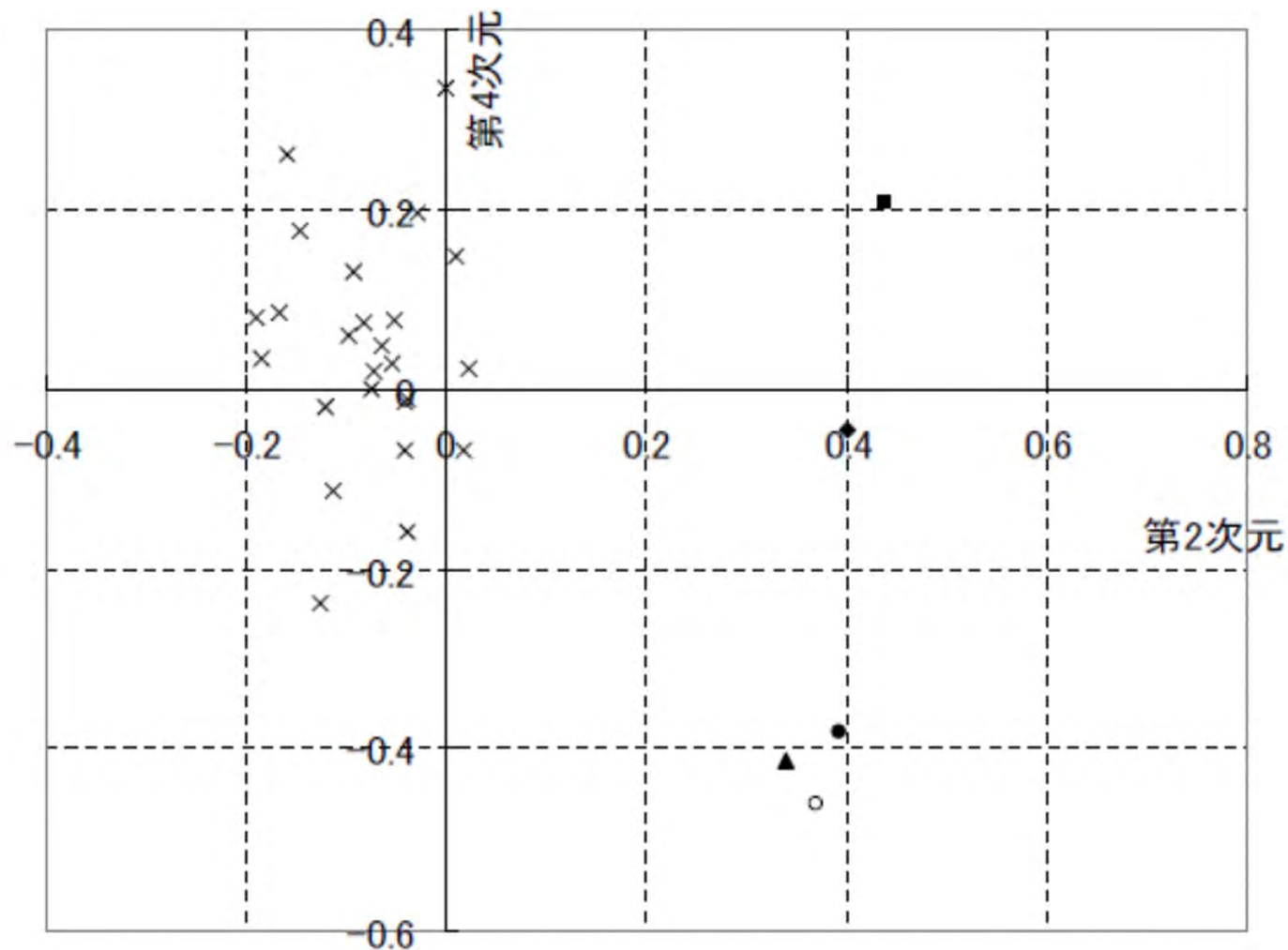


図5 個人差MDSの共通対象付置(第2次元×第4次元)

答える人の違い

調査選択バイアス

Q13. あなたが、今後、調査に協力してもよいとお考えの調査の方法はつぎのどれにあたりますか。
(○はいくつでも)

1. 訪問面接調査 (調査員が調査対象者の自宅を訪問し、面接で行う調査)
2. 留置き調査 (調査員が調査対象者の自宅を訪問して調査票を渡し、調査対象者が自分で回答を記入し、後日、調査員が訪問して調査票を回収する調査)
3. 電話調査 (調査員が調査対象者に電話をして行う調査)
4. 郵送調査 (調査対象者の自宅に調査票を郵送し、調査対象者が自分で回答を記入したあと郵送で返送する調査)
5. インターネット調査 (インターネットを通じて調査の依頼や回答を行う調査)
6. いずれも協力したくない

調査選択バイアス

図表1 協力してもよい調査方法（MA）【2007年調査】

		（％）						
	N	訪問面接調査	訪問留置調査	電話調査	郵送調査	ネット調査	どれもいや	無回答
郵送ランダム	753	5.6	18.9	9.7	84.1	39.7	8.0	1.1
郵送モニター	986	20.1	40.6	32.4	94.7	60.1	3.1	0.5
webモニターA	1000	10.9	30.5	22.0	70.4	96.5	2.0	0.0
webモニターB	1000	15.8	36.8	27.4	80.1	97.1	2.1	0.0
webモニターC	984	8.6	26.4	18.5	62.6	92.9	4.8	0.0

モニター効果

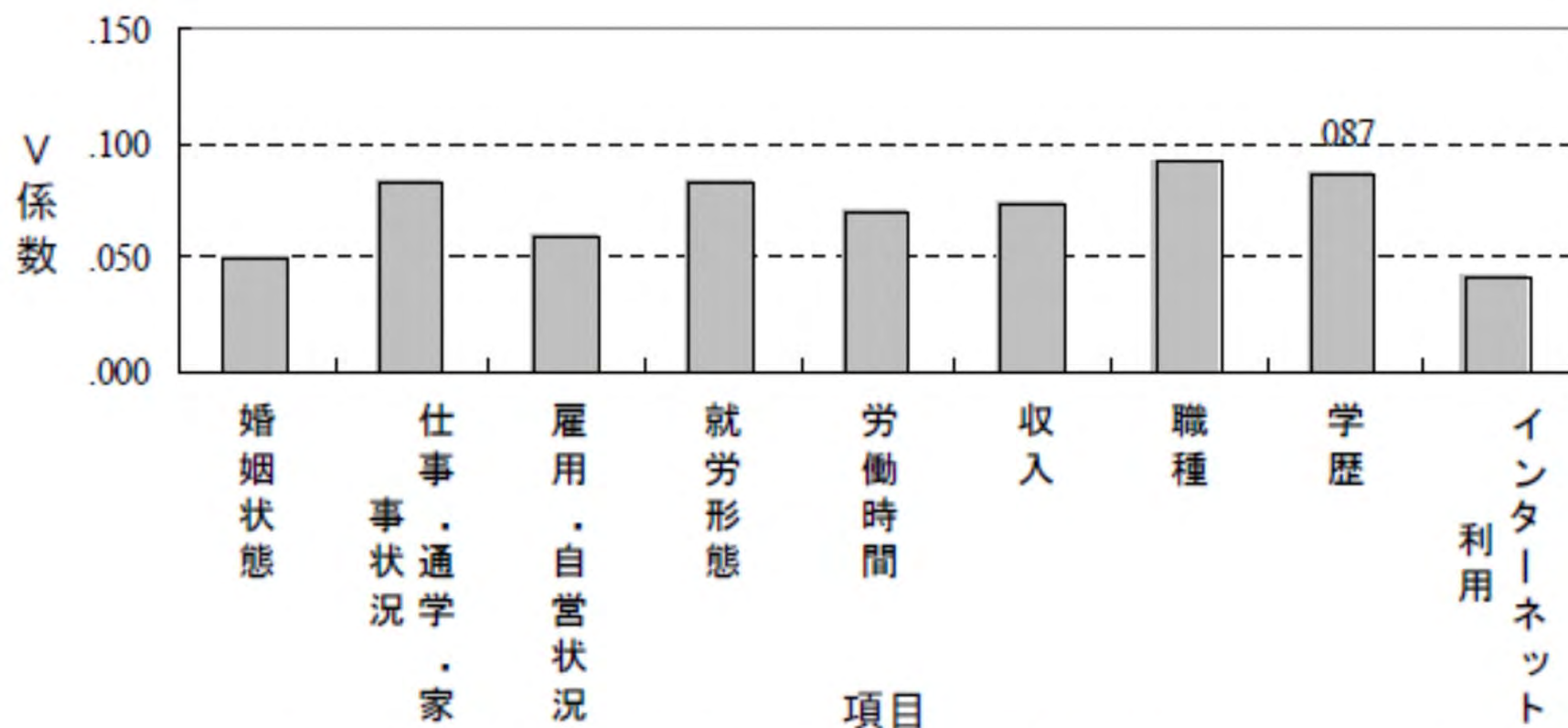


図4 郵送モニターと郵送ランダム属性回答の違い(Cramer's V)

モニター効果

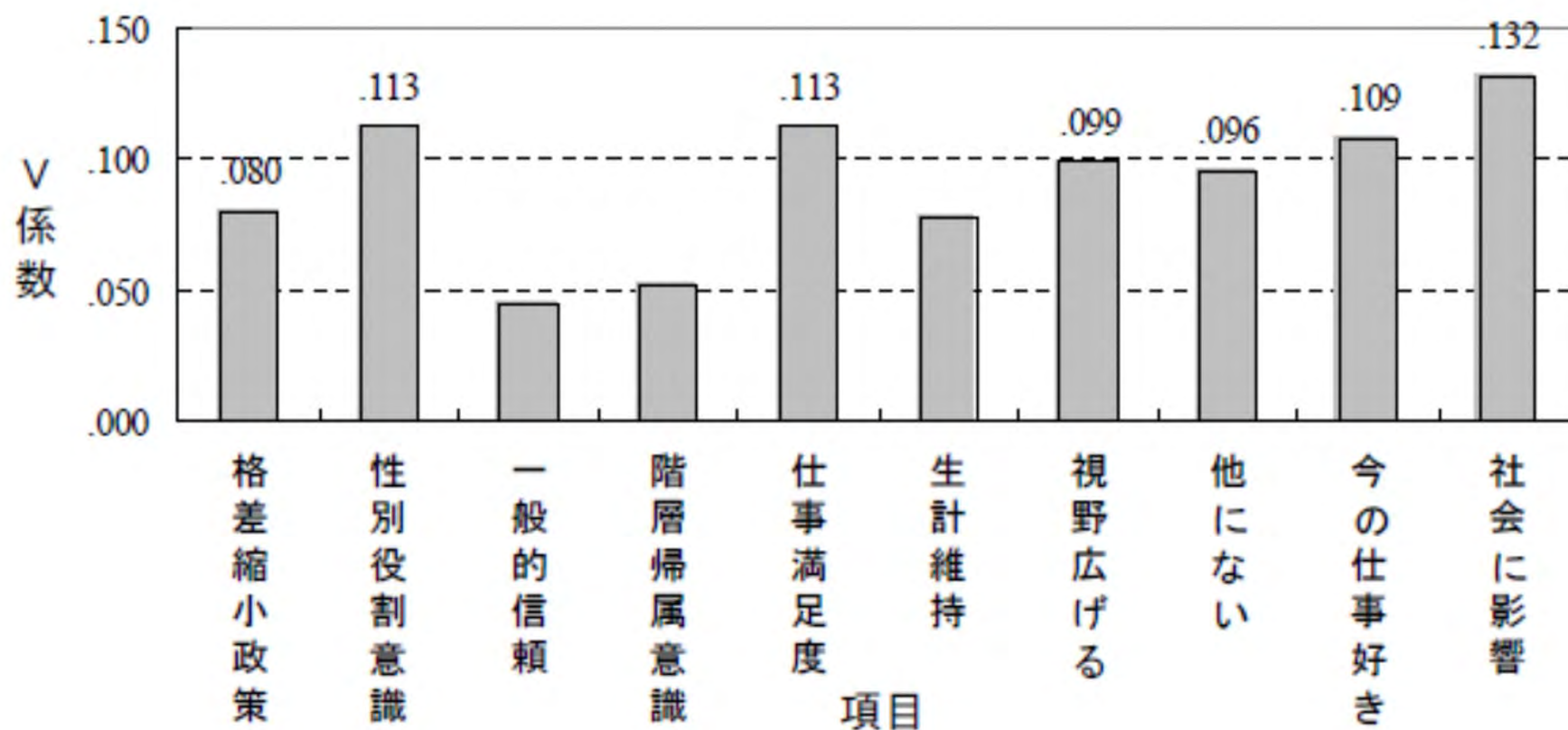


図5 郵送モニターと郵送ランダムを意識回答の違い (Cramer's V)

モニター効果

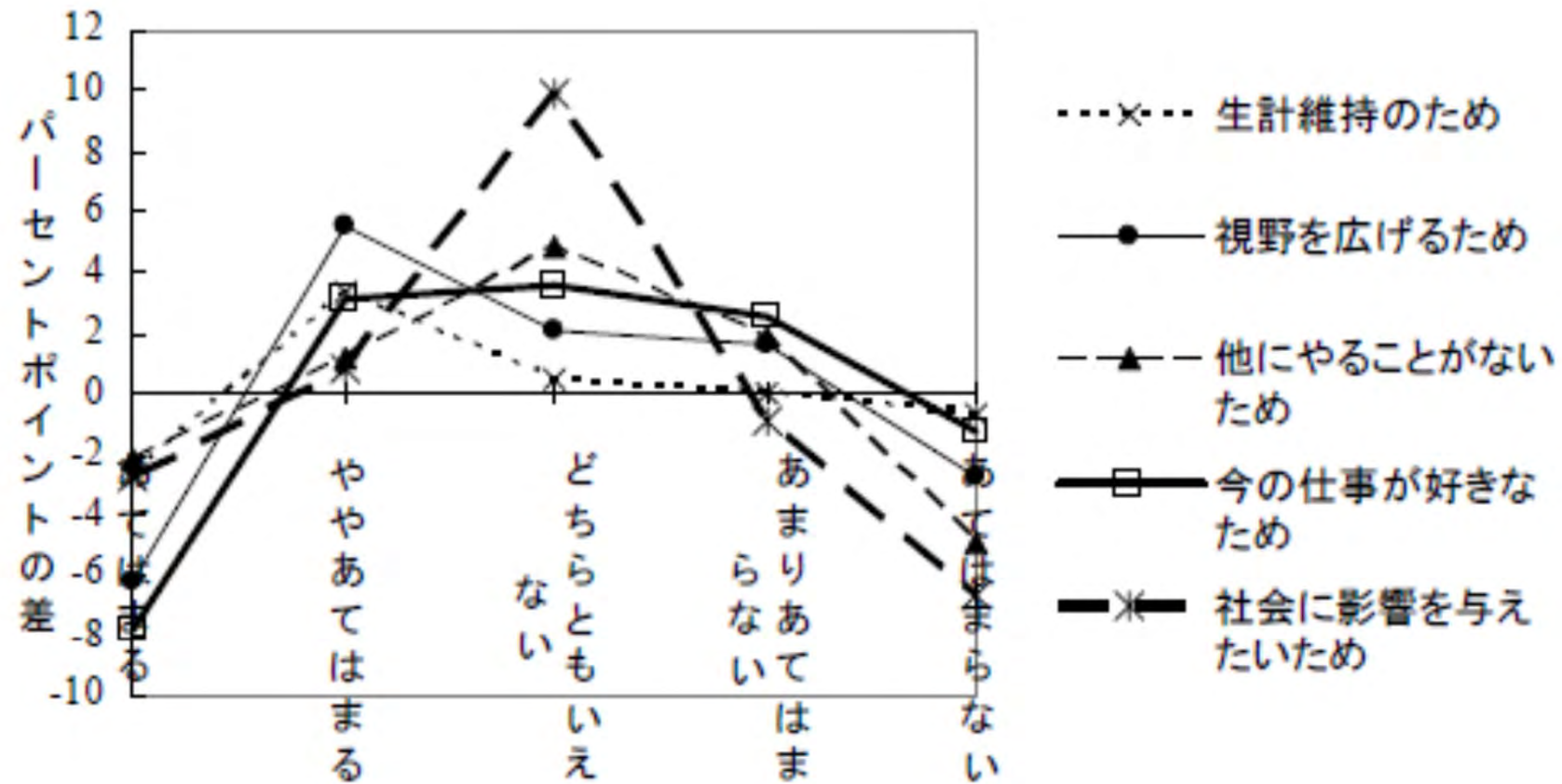


図6 「仕事をする理由」における郵送モニターの特徴
(郵送ランダムを基準とした回答のズレ)

モニター効果



図7 郵送モニターと郵送ランダム意識回答の違い (Cramer's V)

答え方による違い

モード効果

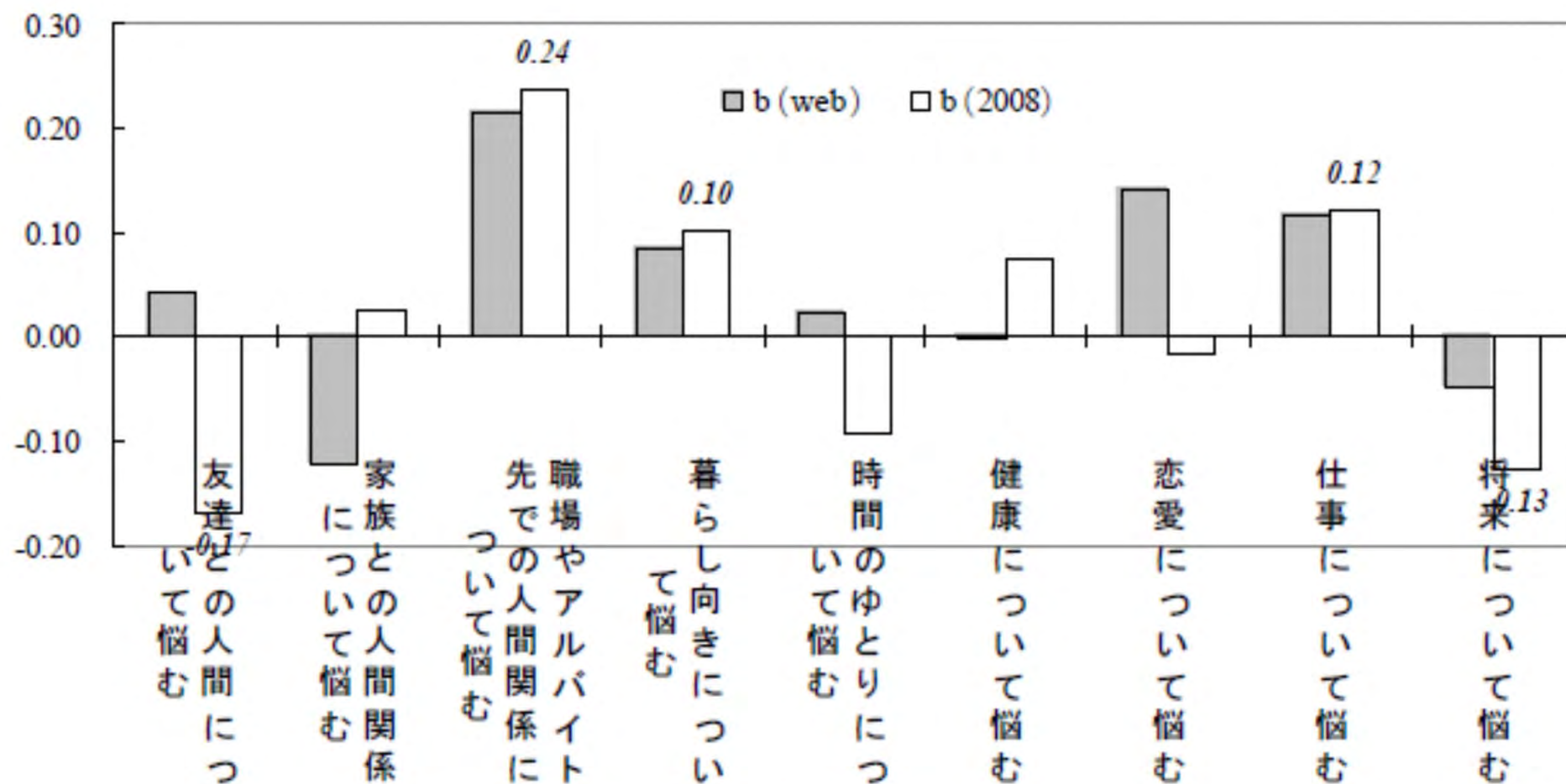


図1 日常生活における悩みに関するモード効果

モード効果

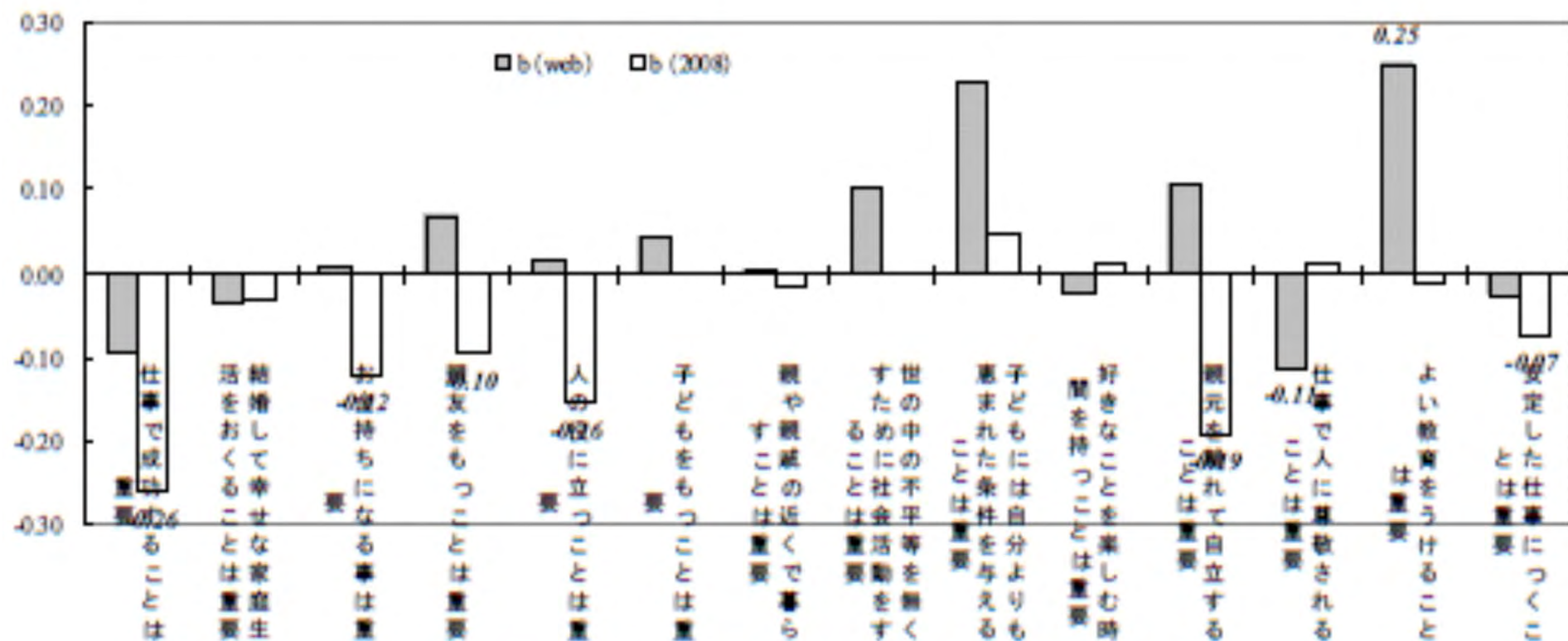


図2 人生にとって重要なものに関するモード効果

モード効果

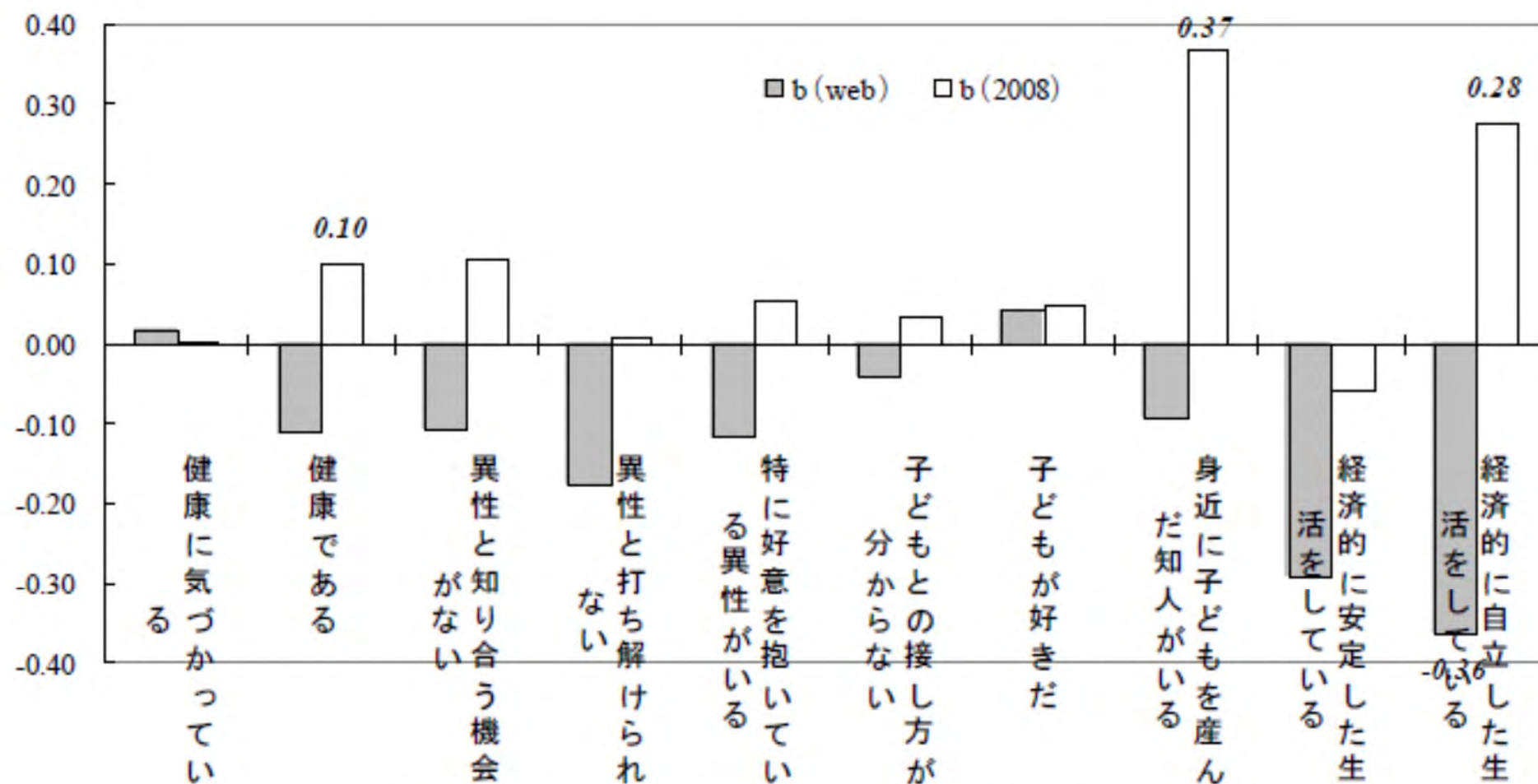


図3 日常生活にかかわる事に関するモード効果

補正の可能性

IPWによる補正事例(三輪 2012)

		(1)	(2)	(3)
		第1波	第3波	第3波
			調整前	調整後
		(N=1,815)	(N=1,439)	(N=1,439)
性別	男性	44.46	42.18	44.56
	女性	55.54	57.82	55.44
年齢	35歳未満	13.94	12.93	13.73
	35歳以上50歳未満	34.82	34.05	34.98
	50歳以上65歳未満	35.04	35.86	35.12
	65歳以上	16.20	17.16	16.17
本人学歴	中学	11.07	9.87	11.19
	高校	44.02	44.61	44.06
	専修学校	11.02	10.35	11.05
	短大・高専	11.13	11.88	10.99
	大学以上	22.75	23.28	22.70
配偶状態	配偶者あり	80.72	83.04	80.85
	無配偶	19.28	16.96	19.15
本人就業	正規就業	48.32	45.93	48.43
	非正規就業	21.21	21.75	21.19
	無職	30.47	32.31	30.38
本人年収	100万円未満	30.25	31.06	30.04
	100万円以上300万円未満	27.99	28.49	28.06
	300万円以上500万円未満	20.17	18.55	20.32
	500万円以上800万円未満	14.88	14.80	14.86
	800万円以上	6.72	7.09	6.73

IPWによる補正事例(三輪 2012)

表3 継続調査実施の段階における「脱落」の要因(プロビット分析)

		全サンプル (N=1,815)		男性 (N=807)		女性 (N=1,008)	
定数項		-1.2986	***	-0.6500		-1.7075	***
性別 [基準:女性]	男性	0.3384	***				
年齢 [基準:65歳以上]	35歳未満	0.3649	**	0.3222		0.4222	*
	35歳以上50歳未満	0.3466	**	0.2413		0.4483	**
	50歳以上65歳未満	0.1453		0.1359		0.1525	
都市規模 [基準:町村]	大都市(18大都市)	-0.0523		0.1581		-0.2634	*
	その他市部	0.0062		-0.0178		0.0073	
地域 [基準:九州・沖縄]	北海道・東北	0.0316		-0.1025		0.1309	
	関東	-0.2464	**	-0.3745	**	-0.1594	
	中部	-0.1612		-0.2698		-0.0901	
	関西	-0.0297		-0.0818		0.0337	
	中国・四国	-0.0313		-0.0390		-0.0209	
配偶状態 [基準:無配偶]	配偶者あり	-0.4494	***	-0.5450	***	-0.3568	**
親との同居 [基準:なし]	同居中	-0.1322		-0.1857		-0.0693	
子どもの有無 [基準:なし]	あり	0.0834		0.0059		0.2459	
未就学児の有無 [基準:なし]	あり	0.0590		0.0004		0.0883	
本人学歴 [基準:大学以上]	中学	0.5065	***	0.4769	***	0.4751	**
	高校	0.1367		0.2662	**	-0.0814	
	専修学校	0.3127	**	0.4077	**	0.1340	
	短大・高専	-0.0124		0.5324	**	-0.3262	*

インターネット調査補正のための 傾向スコア法

- 補正の手順(星野・森本 2007)
 1. 予備的実験調査—従来型とネット—をおこなう
 2. 実験調査データから補正に役立つ変数セットを見つける
 3. 本調査としてのインターネット調査をおこなう
 4. 実験調査データと本調査データを用い、傾向スコアを算出する
 5. 傾向スコアをウェイトとして用いて、本調査データを補正する

インターネット調査補正のための 傾向スコア法

- 今回おこなった代替的手順
 1. 予備的実験調査をおこなう …WPS08
 2. 実験調査のうち半分のケースを用いて、補正に役立つ変数セットを見つける
 3. 本調査の代わりに、残り半分のケースを検証用データとして設定
 4. 実験調査データと検証用データを用い、傾向スコアを算出する
 5. 傾向スコアをウェイトとして用いて、検証用データを補正する

補正の可能性

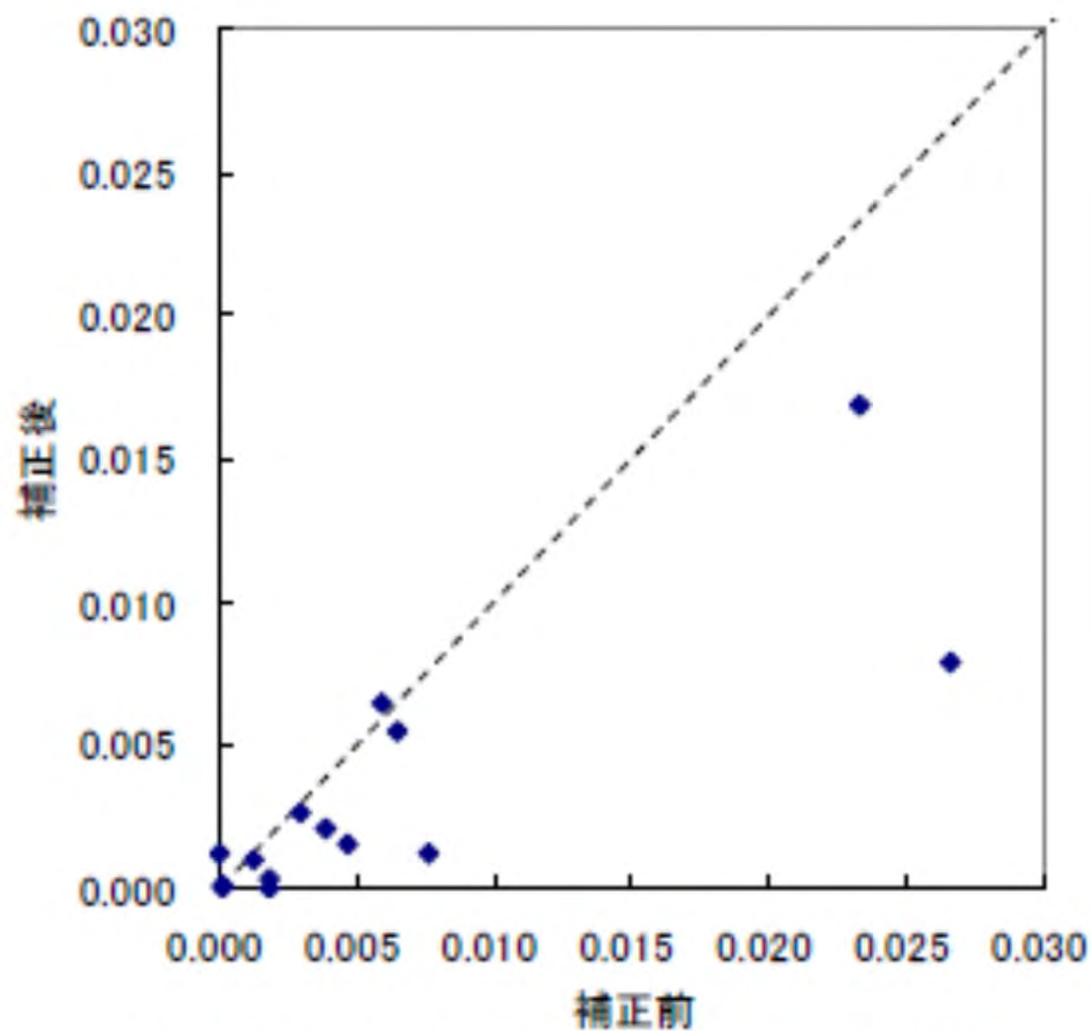


図1 傾向スコア重み付けによる二乗誤差減少

補正の可能性

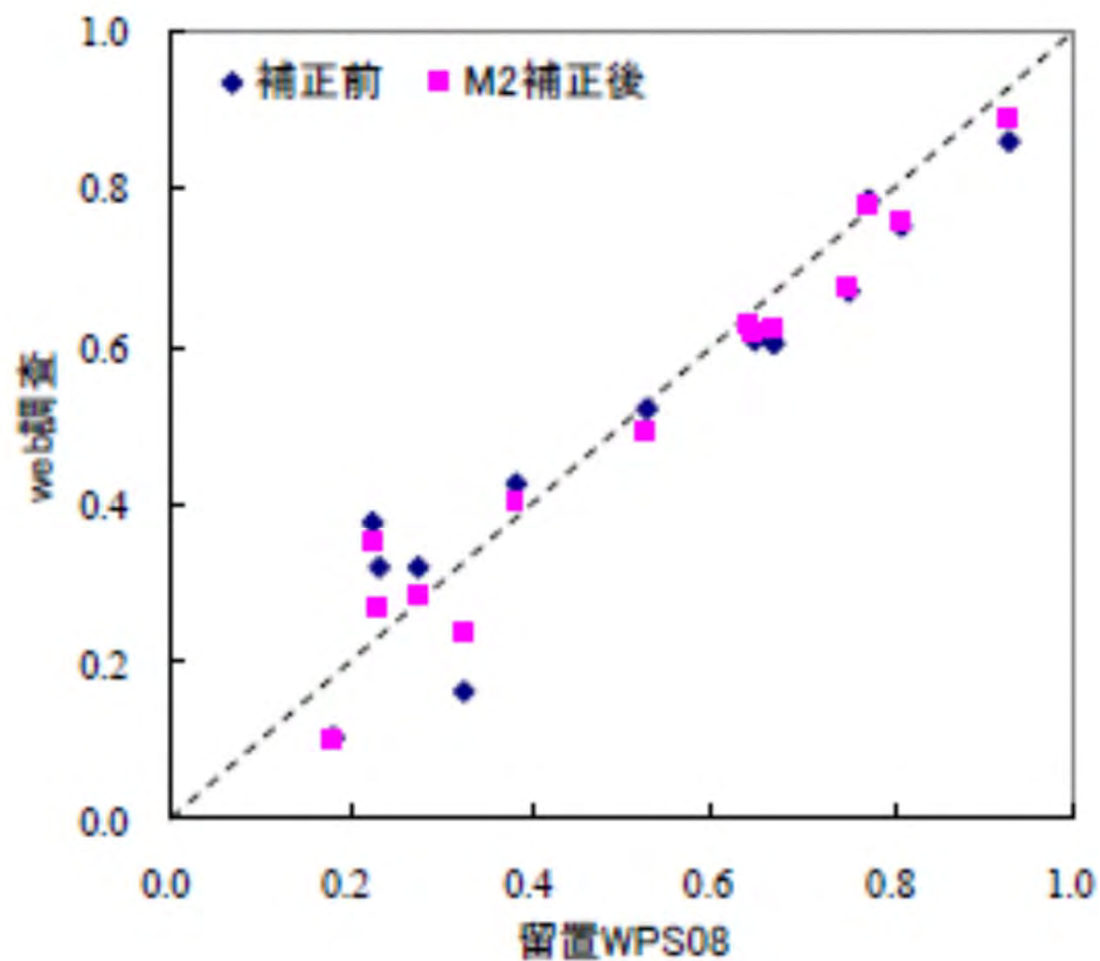


図2 web調査と留置調査の平均比較

補正の可能性

- 傾向スコアによる重み付け補正によって
推定したデータを使った場合は、45%ほど、誤差の
二乗和を減少させた
 - ある程度有効
- ただし、クロスバリデーションの結果、別データでは
減少した割合は、23%にとどまった
 - 物足りない結果

補正の可能性

- 楽観的な話を言うならば・・・
 - もっとよい傾向スコアを作成できたなら、補正はよりよくなったであろう
 - 「調査への志向性」変数など、より役立つものも
- 悲観的な話を言うならば・・・
 - 観察された変数しか、補正には利用できない
 - 関心のある結果変数が絞れるならやりやすいが、一般的に通用する傾向スコアとなると厳しい
 - 補正のためだけの変数を入れる余裕はない

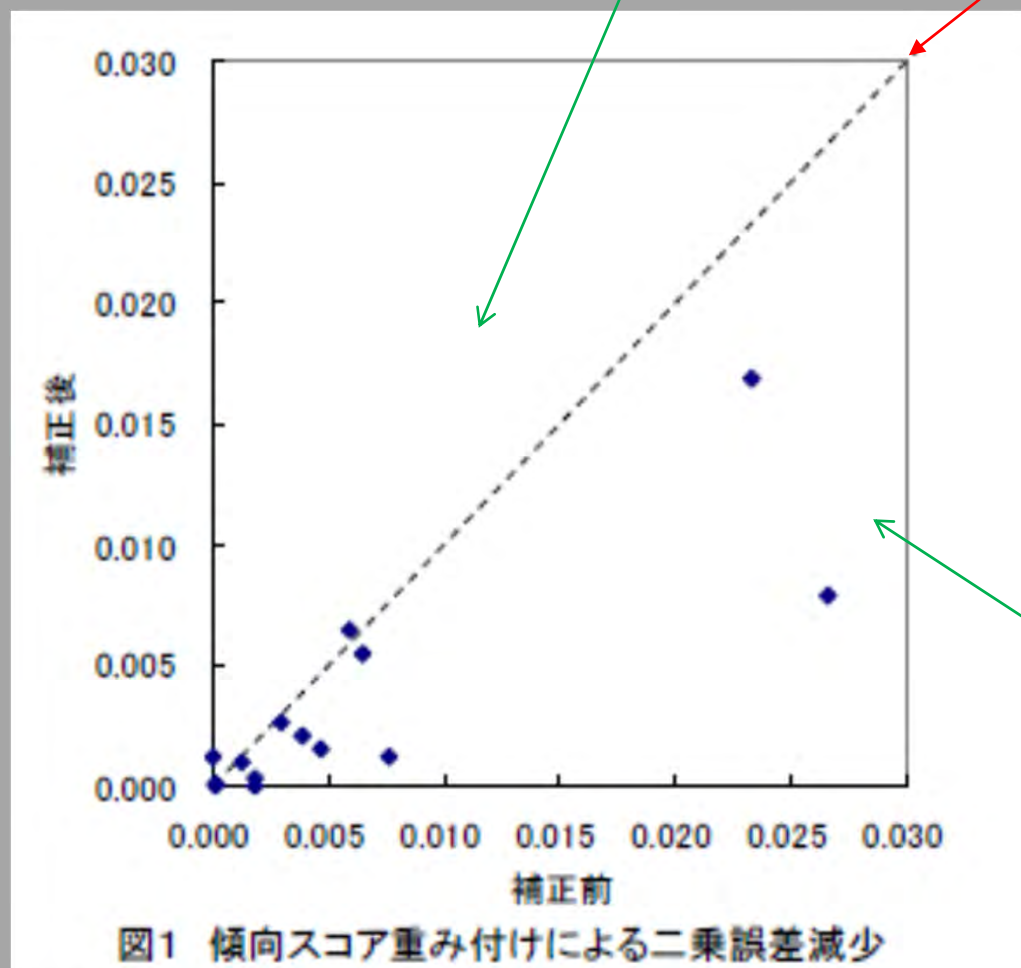
まとめ

- モード効果が必ずしも顕著ではないことから、インターネット調査の偏りは主に「誰が回答するか」の違いと推測
- モニターを使うことで、とりわけ意識項目の分布は大きく変わりうる
- 調査選択バイアスがあるので、従来型調査の回答層をそのまま再現するのは困難
- よって、回答の分布は調査法ごとに大きく異なり、その補正は「特別な質問」を入れない限り限定的

スライド40(補正の可能性) 補足

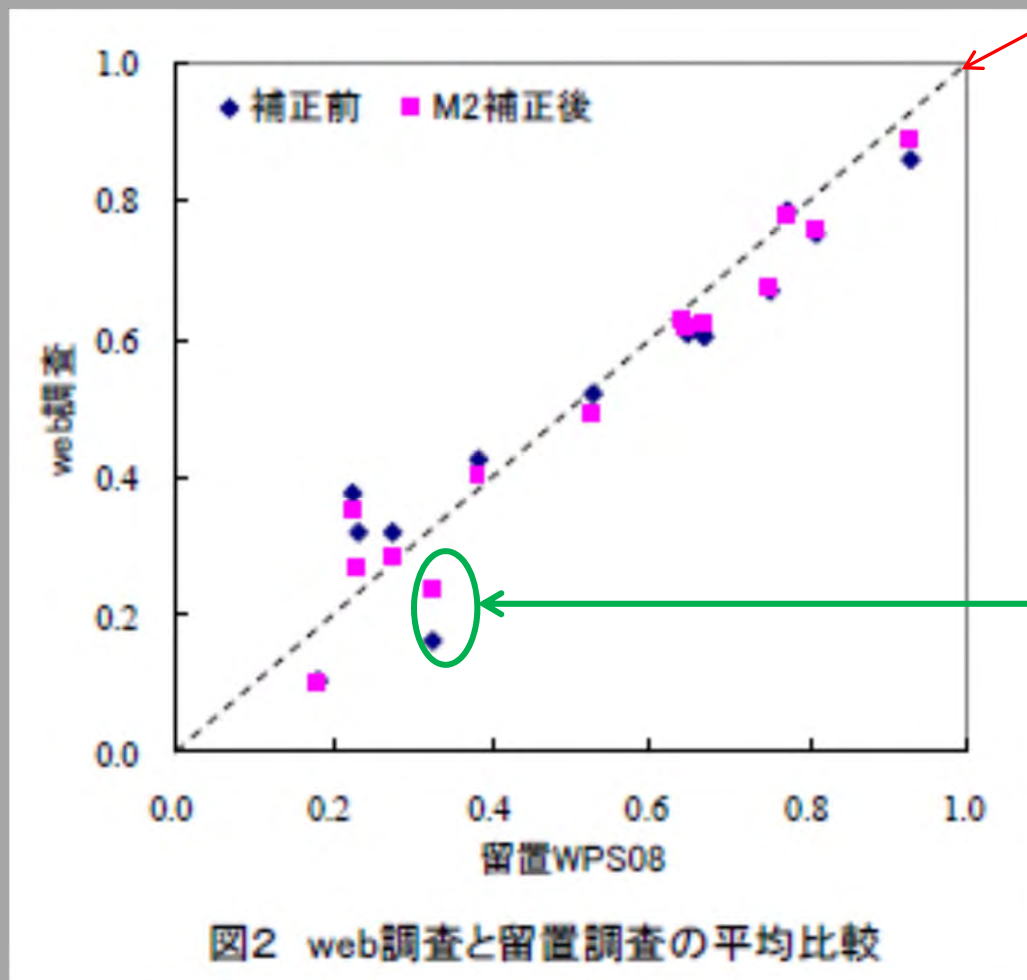
補正の失敗を示す領域
(補正後数値 > 補正前数値は、
補正による誤差の増大を意味する)

補正前と補正後の誤差が同じ
こと (補正による変化なし)
を示すライン



補正の成功を
示す領域

スライド41 (補正の可能性) 補足



留置とWEBの平均値が一致することを意味するライン (45度線)

■が◆よりも、45度線に近づいていれば、補正は成功している。