第6章：パネルデータと差の差分析

Jonathan Roth

数理計量経済学 I
ブラウン大学

モチベーション

- これまで、観察可能な特徴を条件として、処置がランダムに割り当てられていると見なせる（条件付き非交絡性）場合に、因果的効果を推定する方法を見てきました。
- しかし、適切に考慮できていない観察不可能な特徴（すなわち、交絡変数）が存在することを懸念することがよくあります。

モチベーション

- これまで、観察可能な特徴を条件として、処置がランダムに割り当てられていると見なせる（条件付き非交絡性）場合に、因果的効果を推定する方法を見てきました。
- しかし、適切に考慮できていない観察不可能な特徴（すなわち、交絡変数）が存在することを懸念することがよくあります。
- 次に、パネルデータがある場合に、特定の種類の観察されない交絡因子にどのように対処できるかを考えます。

パネルデータとは何か？

- パネルデータとは、各ユニット i（個人や州など）について、複数の期間 t にわたって観測値が得られる状況を指します。

パネルデータとは何か？

- パネルデータとは、各ユニット i（個人や州など）について、複数の期間 t にわたって観測値が得られる状況を指します。
- なぜこれが有用なのでしょうか？ これにより、処置が発生する前の、処置群と対照群のアウトカムの差を調べることができるからです。

パネルデータとは何か？

- パネルデータとは、各ユニット i（個人や州など）について、複数の期間 t にわたって観測値が得られる状況を指します。
- なぜこれが有用なのでしょうか？ これにより、処置が発生する前の、処置群と対照群のアウトカムの差を調べることができるからです。
- もし処置前に処置群と対照群のアウトカムが異なっていれば、それは交絡因子の結果であるはずです。

パネルデータとは何か？

- パネルデータとは、各ユニット i（個人や州など）について、複数の期間 t にわたって観測値が得られる状況を指します。
- なぜこれが有用なのでしょうか？ これにより、処置が発生する前の、処置群と対照群のアウトカムの差を調べることができるからです。
- もし処置前に処置群と対照群のアウトカムが異なっていれば、それは交絡因子の結果であるはずです。
- したがって、処置前の差を利用して交絡因子について学び、それらを調整できる可能性があります。

パネルデータとは何か？

- パネルデータとは、各ユニット i（個人や州など）について、複数の期間 t にわたって観測値が得られる状況を指します。
- なぜこれが有用なのでしょうか？ これにより、処置が発生する前の、処置群と対照群のアウトカムの差を調べることができるからです。
- もし処置前に処置群と対照群のアウトカムが異なっていれば、それは交絡因子の結果であるはずです。
- したがって、処置前の差を利用して交絡因子について学び、それらを調整できる可能性があります。
- 応用マイクロ経済学の研究で最も一般的に使用されるパネルデータ手法である、差の差分析（difference-in-differences, DID）の例で、これがどのように機能するかを見てみましょう。

Hastings (2004)

- 2004 年、Justine Hastings（ブラウン大学の元教授です！）は、ガソリン業界の合併がガソリン価格にどのように影響するかを分析した研究を発表しました。

Hastings (2004)

- 2004 年、Justine Hastings（ブラウン大学の元教授です！）は、ガソリン業界の合併がガソリン価格にどのように影響するかを分析した研究を発表しました。
- 特に彼女は、精製業者である ARCO が、最大手ガソリンスタンドの 1 つである Thrifty を買収したカリフォルニアでの事例を研究しました。

Hastings (2004)

- 2004 年、Justine Hastings（ブラウン大学の元教授です！）は、ガソリン業界の合併がガソリン価格にどのように影響するかを分析した研究を発表しました。
- 特に彼女は、精製業者である ARCO が、最大手ガソリンスタンドの 1 つである Thrifty を買収したカリフォルニアでの事例を研究しました。
- このような合併は価格にどのような影響を与えると思いますか？

Hastings (2004)

- 2004 年、Justine Hastings（ブラウン大学の元教授です！）は、ガソリン業界の合併がガソリン価格にどのように影響するかを分析した研究を発表しました。
- 特に彼女は、精製業者である ARCO が、最大手ガソリンスタンドの 1 つである Thrifty を買収したカリフォルニアでの事例を研究しました。
- このような合併は価格にどのような影響を与えると思いますか？
  - 一方で、競争が減少し、価格が上昇する可能性があります。
  - 他方で、合併によってガソリン供給のコストが削減され、価格が下落する可能性もあります（シナジー効果)。

Hastings (2004)

- 2004 年、Justine Hastings（ブラウン大学の元教授です！）は、ガソリン業界の合併がガソリン価格にどのように影響するかを分析した研究を発表しました。
- 特に彼女は、精製業者である ARCO が、最大手ガソリンスタンドの 1 つである Thrifty を買収したカリフォルニアでの事例を研究しました。
- このような合併は価格にどのような影響を与えると思いますか？
  - 一方で、競争が減少し、価格が上昇する可能性があります。
  - 他方で、合併によってガソリン供給のコストが削減され、価格が下落する可能性もあります（シナジー効果）。
- Hastings は、カリフォルニア州の近隣地域別のガソリン価格データを用いて、この問いに実証的に答えようとしました。
  - データには、Thrifty スタンドがある地域とない地域の両方の情報が含まれています。

- まず、合併後のガソリン価格データしか手に入らないと仮定しましょう。
- 以前に Thrifty があった地域 ($D_i = 1$) と、以前に Thrifty がなかった地域 ($D_i = 0$) の価格を比較することで、Thrifty の転換による因果的効果を推定できるかもしれません。

- まず、合併後のガソリン価格データしか手に入らないと仮定しましょう。

- 以前に Thrifty があった地域 $(D_i = 1)$ と、以前に Thrifty がなかった地域 $(D_i = 0)$ の価格を比較することで、Thrifty の転換による因果的効果を推定できるかもしれません。

- なぜこれが Thrifty 転換の因果的効果を与えない可能性があるのでしょうか？

- まず、合併後のガソリン価格データしか手に入らないと仮定しましょう。

- 以前に Thrifty があった地域 $(D_i = 1)$ と、以前に Thrifty がなかった地域 $(D_i = 0)$ の価格を比較することで、Thrifty の転換による因果的効果を推定できるかもしれません。

- なぜこれが Thrifty 転換の因果的効果を与えない可能性があるのでしょうか？ 欠落変数です！

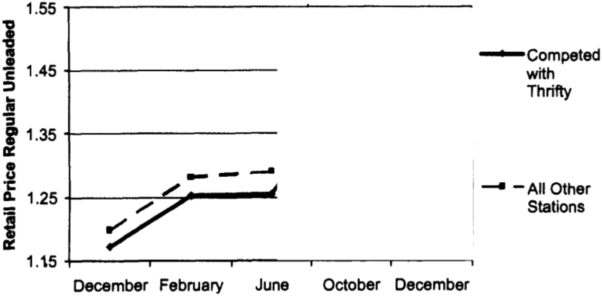
- まず、合併後のガソリン価格データしか手に入らないと仮定しましょう。

- 以前に Thrifty があった地域 $(D_i = 1)$ と、以前に Thrifty がなかった地域 $(D_i = 0)$ の価格を比較することで、Thrifty の転換による因果的効果を推定できるかもしれません。

- なぜこれが Thrifty 転換の因果的効果を与えない可能性があるのでしょうか？ 欠落変数です！

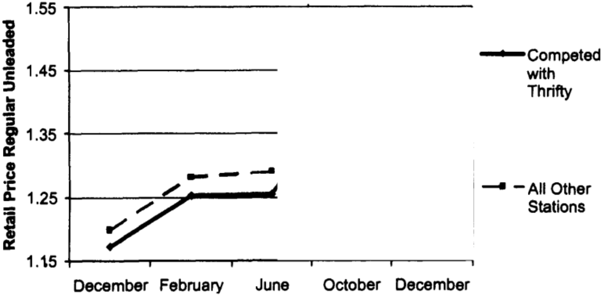
- 特に、以前から Thrifty があった場所は、なかった場所よりも競争が激しかった可能性があります。そのため、もともと価格が低いことが予想されます。

- まず、合併後のガソリン価格データしか手に入らないと仮定しましょう。
- 以前に Thrifty があった地域 ($D_i = 1$) と、以前に Thrifty がなかった地域 ($D_i = 0$) の価格を比較することで、Thrifty の転換による因果的効果を推定できるかもしれません。
- なぜこれが Thrifty 転換の因果的効果を与えない可能性があるのでしょうか？ 欠落変数です！
- 特に、以前から Thrifty があった場所は、なかった場所よりも競争が激しかった可能性があります。そのため、もともと価格が低いことが予想されます。
- パネルデータがあれば、合併前の価格を見ることで、これを実証的にテストできます！



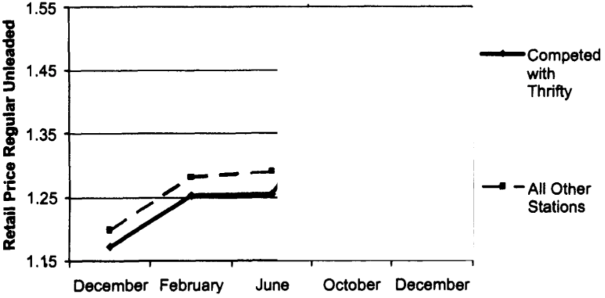
(a) Los Angeles

- 合併前、Thrifty と競合していた市場のスタンドは、どの期間においてもガソリン価格が約 3 セント低かったことがわかります。



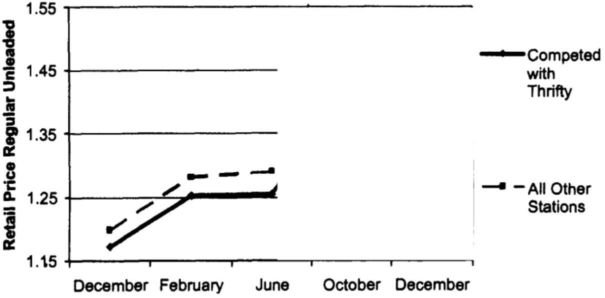
(a) LOS ANGELES

- 合併前、Thrifty と競合していた市場のスタンドは、どの期間においてもガソリン価格が約 3 セント低かったことがわかります。
- 合併後の非交絡性を仮定するのは妥当でしょうか？



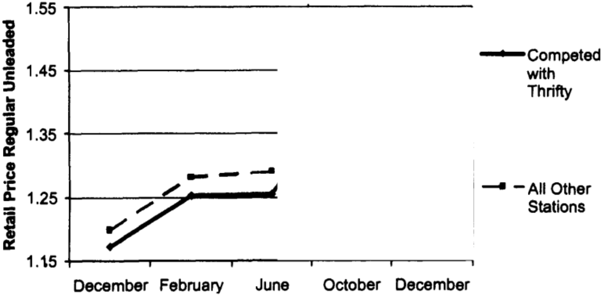
(a) LOS ANGELES

- 合併前、Thrifty と競合していた市場のスタンドは、どの期間においてもガソリン価格が約 3 セント低かったことがわかります。
- 合併後の非交絡性を仮定するのは妥当でしょうか？ いいえ！



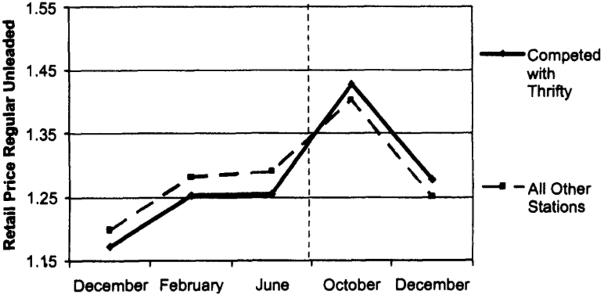
(a) LOS ANGELES

- 合併前、Thrifty と競合していた市場のスタンドは、どの期間においてもガソリン価格が約 3 セント低かったことがわかります。
- 合併後の非交絡性を仮定するのは妥当でしょうか？ いいえ！
- より妥当な仮定は、もし合併がなければ、その 3 セントの差が維持されていたであろう、というものです！



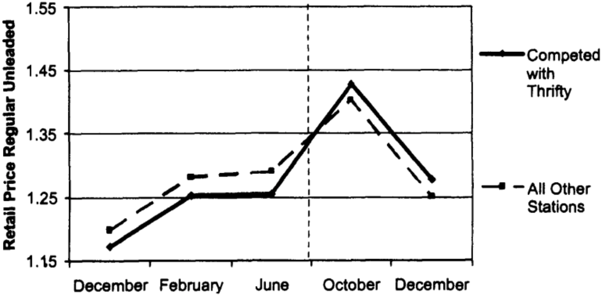
(a) Los Angeles

- 合併前、Thrifty と競合していた市場のスタンドは、どの期間においてもガソリン価格が約 3 セント低かったことがわかります。
- 合併後の非交絡性を仮定するのは妥当でしょうか？ いいえ！
- より妥当な仮定は、もし合併がなければ、その 3 セントの差が維持されていたであろう、というものです！ これが差の差分析の考え方です。



(a) LOS ANGELES

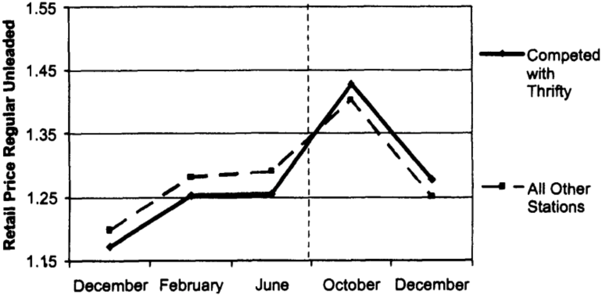
- 合併後、Thrifty があった地域のスタンドは価格が約 2 セント高くなりました。



(a) LOS ANGELES

- 合併後、Thrifty があった地域のスタンドは価格が約 2 セント高くなりました。

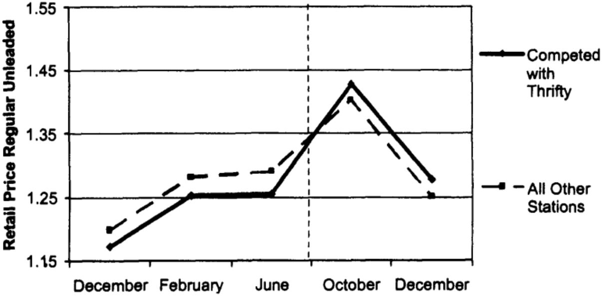
- もし（合併前のように）価格が 3 セント低かったはずだと仮定すると、処置効果は



(a) Los Angeles

- 合併後、Thrifty があった地域のスタンドは価格が約 2 セント高くなりました。

- もし（合併前のように）価格が 3 セント低かったはずだと仮定すると、処置効果は $2-(-3)=5$ セントとなります。



(a) Los Angeles

- 合併後、Thrifty があった地域のスタンドは価格が約 2 セント高くなりました。

- もし（合併前のように）価格が 3 セント低かったはずだと仮定すると、処置効果は $2 - (-3) = 5$ セントとなります。

- これは、処置後の処置群と対照群の差 (2) から、処置前の差 (-3) を引いたもの、すなわち「差の差」です。

DID の仮定の定式化

- 2 期間 $t = 1, 2$ を考えます。処置群 $(D_i = 1)$ は期間 2 で処置を受け、対照群は一度も処置を受けません。
- ユニット i の期間 t における観測されるアウトカムを Y_{it} とします。$Y_{it} = D_i Y_{it}(1) + (1 - D_i) Y_{it}(0)$ と仮定します。

DID の仮定の定式化

- 2 期間 $t = 1, 2$ を考えます。処置群 ($D_i = 1$) は期間 2 で処置を受け、対照群は一度も処置を受けません。
- ユニット i の期間 t における観測されるアウトカムを Y_{it} とします。$Y_{it} = D_i Y_{it}(1) + (1 - D_i) Y_{it}(0)$ と仮定します。
- 事前反応なしの仮定 (No anticipation)：$Y_{i1}(0) = Y_{i1}(1)$
  - 期間 2 での処置が、期間 1 のアウトカムに影響を与えないという仮定です。
- 平行トレンド仮定 (Parallel trends)：

$$\underbrace{E[Y_{i2}(0) - Y_{i1}(0) | D_i = 1]}_{\text{処置群の } Y(0) \text{ の変化}} = \underbrace{E[Y_{i2}(0) - Y_{i1}(0) | D_i = 0]}_{\text{対照群の } Y(0) \text{ の変化}}$$

DID の仮定の定式化

- 2 期間 $t = 1, 2$ を考えます。処置群 ($D_i = 1$) は期間 2 で処置を受け、対照群は一度も処置を受けません。

- ユニット i の期間 t における観測されるアウトカムを Y_{it} とします。$Y_{it} = D_i Y_{it}(1) + (1 - D_i) Y_{it}(0)$ と仮定します。

- 事前反応なしの仮定 (No anticipation)：$Y_{i1}(0) = Y_{i1}(1)$
  - 期間 2 での処置が、期間 1 のアウトカムに影響を与えないという仮定です。

- 平行トレンド仮定 (Parallel trends)：

$$\underbrace{E[Y_{i2}(0) - Y_{i1}(0) | D_i = 1]}_{\text{処置群の } Y(0) \text{ の変化}} = \underbrace{E[Y_{i2}(0) - Y_{i1}(0) | D_i = 0]}_{\text{対照群の } Y(0) \text{ の変化}}$$

  同値な表現として：

$$\underbrace{E[Y_{i2}(0) | D_i = 1] - E[Y_{i2}(0) | D_i = 0]}_{\text{期間 2 における選択バイアス}} = \underbrace{E[Y_{i1}(0) | D_i = 1] - E[Y_{i1}(0) | D_i = 0]}_{\text{期間 1 における選択バイアス}}$$

- これらの仮定の下で、以下が成り立ちます：

$$\underbrace{E[Y_{i2} - Y_{i1}|D_i = 1]}_{\text{処置群の観測された変化}} - \underbrace{E[Y_{i2} - Y_{i1}|D_i = 0]}_{\text{対照群の観測された変化}} =$$

$$= E[Y_{i2}(1) - Y_{i1}(1)|D_i = 1] - E[Y_{i2}(0) - Y_{i1}(0)|D_i = 0] \text{ (観測データ規則)}$$

- これらの仮定の下で、以下が成り立ちます：

$$
\begin{aligned}
&\underbrace{E[Y_{i2}-Y_{i1}|D_i=1]}_{\text{処置群の観測された変化}}-\underbrace{E[Y_{i2}-Y_{i1}|D_i=0]}_{\text{対照群の観測された変化}}= \\
&=E[Y_{i2}(1)-Y_{i1}(1)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (観測データ規則)} \\
&=E[Y_{i2}(1)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (事前反応なし)}
\end{aligned}
$$

- これらの仮定の下で、以下が成り立ちます：

$$\underbrace{E[Y_{i2}-Y_{i1}|D_i=1]}_{\text{処置群の観測された変化}}-\underbrace{E[Y_{i2}-Y_{i1}|D_i=0]}_{\text{対照群の観測された変化}}=$$

$$=E[Y_{i2}(1)-Y_{i1}(1)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (観測データ規則)}$$

$$=E[Y_{i2}(1)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (事前反応なし)}$$

$$=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1]+$$

$$E[Y_{i2}(0)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{(項の加減)}$$

- これらの仮定の下で、以下が成り立ちます：

$$\underbrace{E[Y_{i2}-Y_{i1}|D_i=1]}_{\text{処置群の観測された変化}}-\underbrace{E[Y_{i2}-Y_{i1}|D_i=0]}_{\text{対照群の観測された変化}}=$$

$$=E[Y_{i2}(1)-Y_{i1}(1)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (観測データ規則)}$$

$$=E[Y_{i2}(1)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (事前反応なし)}$$

$$=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1]+$$

$$E[Y_{i2}(0)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{(項の加減)}$$

$$=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1] \text{ (平行トレンド)}$$

- これらの仮定の下で、以下が成り立ちます：

$$\underbrace{E[Y_{i2}-Y_{i1}|D_i=1]}_{\text{処置群の観測された変化}}-\underbrace{E[Y_{i2}-Y_{i1}|D_i=0]}_{\text{対照群の観測された変化}}=$$

$$=E[Y_{i2}(1)-Y_{i1}(1)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (観測データ規則)}$$

$$=E[Y_{i2}(1)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0] \text{ (事前反応なし)}$$

$$=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1]+$$

$$E[Y_{i2}(0)-Y_{i1}(0)|D_i=1]-E[Y_{i2}(0)-Y_{i1}(0)|D_i=0]\text{(項の加減)}$$

$$=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1] \text{ (平行トレンド)}$$

- したがって、標本平均の「差の差」は $\tau_{ATT}=E[Y_{i2}(1)-Y_{i2}(0)|D_i=1]$ を識別します。

- これは処置群における平均処置効果 (Average Treatment Effect on the Treated, ATT) と呼ばれます。
処置群の、期間 2 における平均的な効果です。

ATT の推定

- DID の仮定（平行トレンドと事前反応なし）の下で、ATT は次のように識別されることを示しました：

$$\tau_{ATT} = \underbrace{E[Y_{i2} - Y_{i1} | D_i = 1]}_{\text{処置群の母平均の変化}} - \underbrace{E[Y_{i2} - Y_{i1} | D_i = 0]}_{\text{対照群の母平均の変化}}$$

- これをどう推定すればよいでしょうか？

ATT の推定

- DID の仮定（平行トレンドと事前反応なし）の下で、ATT は次のように識別されることを示しました：

$$\tau_{ATT} = \underbrace{E[Y_{i2} - Y_{i1}|D_i = 1]}_{\text{処置群の母平均の変化}} - \underbrace{E[Y_{i2} - Y_{i1}|D_i = 0]}_{\text{対照群の母平均の変化}}$$

- これをどう推定すればよいでしょうか？　標本平均を代入すればよいのです！

ATT の推定

- DID の仮定（平行トレンドと事前反応なし）の下で、ATT は次のように識別されることを示しました：

$$\tau_{ATT} = \underbrace{E[Y_{i2} - Y_{i1}|D_i = 1]}_{\text{処置群の母平均の変化}} - \underbrace{E[Y_{i2} - Y_{i1}|D_i = 0]}_{\text{対照群の母平均の変化}}$$

- これをどう推定すればよいでしょうか？　標本平均を代入すればよいのです！

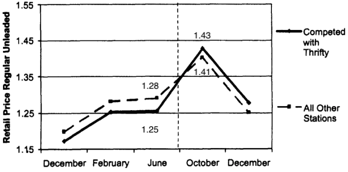
- 推定値は：

$$\hat{\tau}_{ATT} = \underbrace{\bar{Y}_{12} - \bar{Y}_{11}}_{\text{処置群の標本平均の変化}} - \underbrace{\bar{Y}_{02} - \bar{Y}_{01}}_{\text{対照群の標本平均の変化}},$$

ここで $\bar{Y}_{dt}$ は、期間 t における処置ステータス $D_i = d$ のユニットの標本平均です。

例題

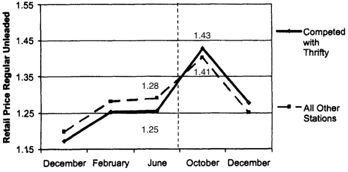
- Hastings の例で、6 月（期間 1）と 10 月を比較してみましょう。



(a) LOS ANGELES

例題

- Hastings の例で、6 月（期間 1）と 10 月を比較してみましょう。



(a) Los Angeles

$$\hat{\tau}_{ATT} = \underbrace{\bar{Y}_{12} - \bar{Y}_{11}}_{\text{処置群の変化}} - \underbrace{\bar{Y}_{02} - \bar{Y}_{01}}_{\text{対照群の変化}} = (1.43 - 1.25) - (1.41 - 1.28) = 0.05$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

ここで $Post_t = 1[t = 2]$ です。

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$E[Y_{it}|D_i = 0, Post_t = 0] =$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$E[Y_{it} | D_i = 0, Post_t = 0] = \beta_0$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$E[Y_{it}|D_i = 0, Post_t = 0] = \beta_0$$
$$E[Y_{it}|D_i = 0, Post_t = 1] =$$

回帰としての DID

- 以下の回帰を考えます：
$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$
ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：
$$E[Y_{it}|D_i = 0, Post_t = 0] = \beta_0$$
$$E[Y_{it}|D_i = 0, Post_t = 1] = \beta_0 + \beta_1$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$\begin{aligned}
E[Y_{it}|D_i = 0, Post_t = 0] &= \beta_0 \\
E[Y_{it}|D_i = 0, Post_t = 1] &= \beta_0 + \beta_1 \\
E[Y_{it}|D_i = 1, Post_t = 0] &=
\end{aligned}$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$E[Y_{it}|D_i = 0, Post_t = 0] = \beta_0$$
$$E[Y_{it}|D_i = 0, Post_t = 1] = \beta_0 + \beta_1$$
$$E[Y_{it}|D_i = 1, Post_t = 0] = \beta_0 + \beta_2$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$E[Y_{it}|D_i = 0, Post_t = 0] = \beta_0$$
$$E[Y_{it}|D_i = 0, Post_t = 1] = \beta_0 + \beta_1$$
$$E[Y_{it}|D_i = 1, Post_t = 0] = \beta_0 + \beta_2$$
$$E[Y_{it}|D_i = 1, Post_t = 1] =$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$\begin{aligned}
E[Y_{it}|D_i = 0, Post_t = 0] &= \beta_0 \\
E[Y_{it}|D_i = 0, Post_t = 1] &= \beta_0 + \beta_1 \\
E[Y_{it}|D_i = 1, Post_t = 0] &= \beta_0 + \beta_2 \\
E[Y_{it}|D_i = 1, Post_t = 1] &= \beta_0 + \beta_1 + \beta_2 + \beta_3
\end{aligned}$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

$$\begin{aligned}
E[Y_{it}|D_i = 0, Post_t = 0] &= \beta_0 \\
E[Y_{it}|D_i = 0, Post_t = 1] &= \beta_0 + \beta_1 \\
E[Y_{it}|D_i = 1, Post_t = 0] &= \beta_0 + \beta_2 \\
E[Y_{it}|D_i = 1, Post_t = 1] &= \beta_0 + \beta_1 + \beta_2 + \beta_3
\end{aligned}$$

- したがって、

$$\begin{aligned}
\beta_3 =& (E[Y_{it}|D_i = 1, Post_t = 1] - E[Y_{it}|D_i = 1, Post_t = 0]) - \\
& (E[Y_{it}|D_i = 0, Post_t = 1] - E[Y_{it}|D_i = 0, Post_t = 0]) = \tau_{ATT}
\end{aligned}$$

回帰としての DID

- 以下の回帰を考えます：

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  ここで $Post_t = 1[t = 2]$ です。
- 主張：DID の仮定の下で、母集団回帰係数 β_3 は τ_{ATT} に等しくなります。
- なぜでしょうか？ 上記の回帰は CEF を次のようにモデル化しています：

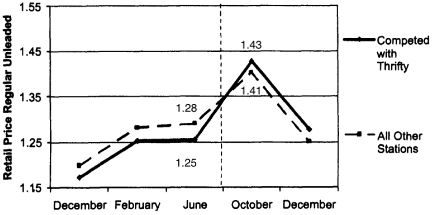
$$\begin{aligned}
E[Y_{it}|D_i = 0, Post_t = 0] &= \beta_0 \\
E[Y_{it}|D_i = 0, Post_t = 1] &= \beta_0 + \beta_1 \\
E[Y_{it}|D_i = 1, Post_t = 0] &= \beta_0 + \beta_2 \\
E[Y_{it}|D_i = 1, Post_t = 1] &= \beta_0 + \beta_1 + \beta_2 + \beta_3
\end{aligned}$$

- したがって、

$$\begin{aligned}
\beta_3 =& (E[Y_{it}|D_i = 1, Post_t = 1] - E[Y_{it}|D_i = 1, Post_t = 0]) - \\
& (E[Y_{it}|D_i = 0, Post_t = 1] - E[Y_{it}|D_i = 0, Post_t = 0]) = \tau_{ATT}
\end{aligned}$$

- 同様に、$\hat{\beta}_3 = (\bar{Y}_{12} - \bar{Y}_{11}) - (\bar{Y}_{02} - \bar{Y}_{01}) = \hat{\tau}_{ATT}$ となります。

例題



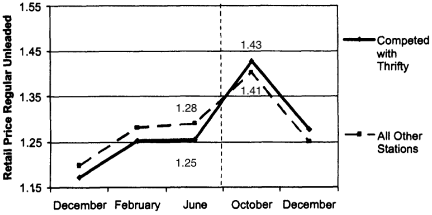
(a) Los Angeles

- Hastings のデータの 6 月と 10 月の部分を使い、

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

を OLS で推定するとします（ $Post_t$ は 10 月なら 1、6 月なら 0)。

例題



(a) LOS ANGELES

- Hastings のデータの 6 月と 10 月の部分を使い、

$$Y_{it} = \beta_0 + \beta_1 \times Post_t + \beta_2 D_i + \beta_3 D_i \times Post_t + \varepsilon_{it},$$

  を OLS で推定するとします（ $Post_t$ は 10 月なら 1、6 月なら 0)。

- 以下の回帰係数が得られます：

定数項 ($\hat{\beta}_0$)	1.28
Post ($\hat{\beta}_1$)	0.13
処置群ダミー ($\hat{\beta}_2$)	-0.03
処置群 $\times$ Post ($\hat{\beta}_3$)	0.05

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。
- これは主に 2 つの理由で有用です：

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。
- これは主に 2 つの理由で有用です：
  1. 処置前に平行トレンドが成り立っているように見えるかをテストできます。

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。
- これは主に 2 つの理由で有用です：
  1. 処置前に平行トレンドが成り立っているように見えるかをテストできます。

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。
- これは主に 2 つの理由で有用です：
  1. 処置前に平行トレンドが成り立っているように見えるかをテストできます。
  2. 時間の経過とともに ATT がどのように変化するかを分析できます。

多期間における DID

- DID 分析において、3 期間以上存在することがよくあります。
- これは主に 2 つの理由で有用です：
  1. 処置前に平行トレンドが成り立っているように見えるかをテストできます。
  2. 時間の経過とともに ATT がどのように変化するかを分析できます。
- これをどのように行うのでしょうか？

多期間における DID

- 期間が $t = -\underline{T}, ..., \bar{T}$ までであり、処置群が期間 1 から処置を受け始めるとします。

多期間における DID

- 期間が $t = -\underline{T}, ..., \bar{T}$ までであり、処置群が期間 1 から処置を受け始めるとします。

- $s \neq 0$ である各期間について、期間 s と期間 0 の間の 2 期間 DID を推定できます：

$$\hat{\beta}_s = \underbrace{(\bar{Y}_{1s} - \bar{Y}_{0s})}_{\text{期間 } s \text{ における差}} - \underbrace{(\bar{Y}_{10} - \bar{Y}_{00})}_{\text{期間 0 における差}}$$

ここで $\bar{Y}_{dt}$ は、期間 t における処置グループ d の平均です。

多期間における DID

- 期間が $t = -\underline{T}, \ldots, \bar{T}$ までであり、処置群が期間 1 から処置を受け始めるとします。

- $s \neq 0$ である各期間について、期間 s と期間 0 の間の 2 期間 DID を推定できます：

$$\hat{\beta}_s = \underbrace{(\bar{Y}_{1s} - \bar{Y}_{0s})}_{\text{期間 } s \text{ における差}} - \underbrace{(\bar{Y}_{10} - \bar{Y}_{00})}_{\text{期間 0 における差}}$$

  ここで $\bar{Y}_{dt}$ は、期間 t における処置グループ d の平均です。

- 便利なことに、これらの $\hat{\beta}_s$ は以下の回帰の OLS 推定値と等しくなります：

$$Y_{it} = \phi_t + D_i\gamma + \sum_{s\neq 0} D_i \times 1[t = s] \times \beta_s + \varepsilon_{it}$$

多期間における DID

- 期間が $t = -\underline{T}, \ldots, \bar{T}$ までであり、処置群が期間 1 から処置を受け始めるとします。

- $s \neq 0$ である各期間について、期間 s と期間 0 の間の 2 期間 DID を推定できます：

$$\hat{\beta}_s = \underbrace{(\bar{Y}_{1s} - \bar{Y}_{0s})}_{\text{期間 } s \text{ における差}} - \underbrace{(\bar{Y}_{10} - \bar{Y}_{00})}_{\text{期間 0 における差}}$$

  ここで $\bar{Y}_{dt}$ は、期間 t における処置グループ d の平均です。

- 便利なことに、これらの $\hat{\beta}_s$ は以下の回帰の OLS 推定値と等しくなります：

$$Y_{it} = \phi_t + D_i\gamma + \sum_{s \neq 0} D_i \times 1[t = s] \times \beta_s + \varepsilon_{it}$$

- $D_i\gamma$ をユニット固定効果 λ_i に置き換えても、全く同じ $\hat{\beta}_s$ が得られます。

例：Medicaid（公的医療保険）の拡大

- 医療費負担適正化法（ACA、通称オバマケア）は、連邦貧困線の 138% までの所得の人々に Medicaid の対象を拡大しました。
- Medicaid の拡大は 2014 年に施行されました。しかし、いくつかの共和党寄りの州は拡大を拒否しました。
- 2015 年までに、24 州が Medicaid を拡大しました（その後、さらに多くの州が拡大しています）。

例：Medicaid（公的医療保険）の拡大

- 医療費負担適正化法（ACA、通称オバマケア）は、連邦貧困線の 138% までの所得の人々に Medicaid の対象を拡大しました。
- Medicaid の拡大は 2014 年に施行されました。しかし、いくつかの共和党寄りの州は拡大を拒否しました。
- 2015 年までに、24 州が Medicaid を拡大しました（その後、さらに多くの州が拡大しています）。
- Carey, Miller, and Wherry (2020) は、早期導入州と非導入州を比較する DID デザインを用いて、Medicaid 拡大の影響を研究しています。

例：Medicaid の拡大

- 彼らの回帰式を簡略化したものは以下の通りです：

$$Y_{its} = \phi_t + \lambda_s + \sum_{r \neq -1} D_i \times 1[t = 2014 + r] \times \beta_r + \varepsilon_{it}$$

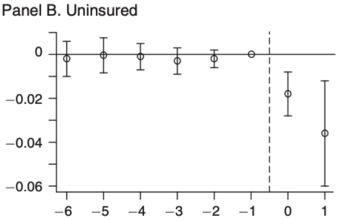
Y_{its} は州 s 年 t の個人 i の結果、$D_i = 1$ は拡大州。β_s の推定値と 95% 信頼区間：

例：Medicaid の拡大

- 彼らの回帰式を簡略化したものは以下の通りです：

$$Y_{its} = \phi_t + \lambda_s + \sum_{r \neq -1} D_i \times 1[t = 2014 + r] \times \beta_r + \varepsilon_{it}$$

Y_{its} は州 s 年 t の個人 i の結果、$D_i = 1$ は拡大州。β_s の推定値と 95% 信頼区間：

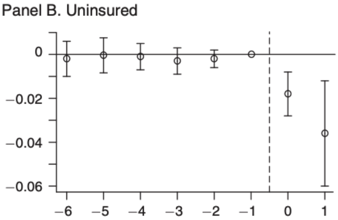


例：Medicaid の拡大

- 彼らの回帰式を簡略化したものは以下の通りです：

$$Y_{its} = \phi_t + \lambda_s + \sum_{r \neq -1} D_i \times 1[t = 2014 + r] \times \beta_r + \varepsilon_{it}$$

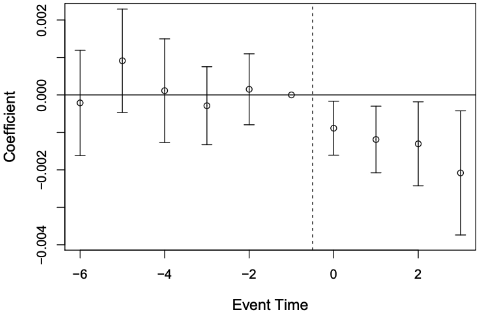
Y_{its} は州 s 年 t の個人 i の結果、$D_i = 1$ は拡大州。β_s の推定値と 95% 信頼区間：



- 処置前は同様のトレンド（「プレトレンド」）ですが、処置後は負の効果が見られます。

関連する論文で、同じ著者の一部が同様のリサーチデザインを用いて、死亡率への影響を推定しています。

Figure 2: Effect of the ACA Medicaid Expansions on Annual Mortality



平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。

平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。これは、時間とともに変化する交絡因子の存在を排除しています。

平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。これは、時間とともに変化する交絡因子の存在を排除しています。
- 私たちはしばしば、時間とともに変化する交絡因子を懸念します。例えば、マクロ経済的要因が民主党支持州と共和党支持州に異なる影響を与えるかもしれません。

平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。これは、時間とともに変化する交絡因子の存在を排除しています。
- 私たちはしばしば、時間とともに変化する交絡因子を懸念します。例えば、マクロ経済的要因が民主党支持州と共和党支持州に異なる影響を与えるかもしれません。
- 処置前の差（「プレトレンド」）をテストすることは、リサーチデザインに対する信頼を高めるのに役立ちます。

平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。これは、時間とともに変化する交絡因子の存在を排除しています。

- 私たちはしばしば、時間とともに変化する交絡因子を懸念します。例えば、マクロ経済的要因が民主党支持州と共和党支持州に異なる影響を与えるかもしれません。

- 処置前の差（「プレトレンド」）をテストすることは、リサーチデザインに対する信頼を高めるのに役立ちます。しかし、それは完璧ではありません。なぜでしょうか？
    1. 以前にトレンドが並行だったからといって、その後も並行であり続けたとは限らないからです。

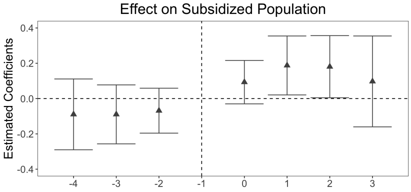
平行トレンドに関する注意点

- DID は平行トレンド仮定に依存しています。これは選択バイアスを許容しますが、それが時間を通じて一定であることを要求します。これは、時間とともに変化する交絡因子の存在を排除しています。
- 私たちはしばしば、時間とともに変化する交絡因子を懸念します。例えば、マクロ経済的要因が民主党支持州と共和党支持州に異なる影響を与えるかもしれません。
- 処置前の差（「プレトレンド」）をテストすることは、リサーチデザインに対する信頼を高めるのに役立ちます。しかし、それは完璧ではありません。なぜでしょうか？
    1. 以前にトレンドが並行だったからといって、その後も並行であり続けたとは限らないからです。
    2. 多くの場合、プレトレンドの推定値にはノイズが含まれるため、本当に 0 なのかどうか確信が持てないからです。

- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。

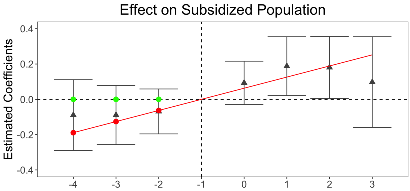
- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。
- プロットが説得力を持つかの目安は、全 CI を通る滑らかな線が引けるかです。

- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。
- プロットが説得力を持つかの目安は、全 CI を通る滑らかな線が引けるかです。



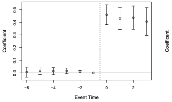
- ここで効果があるという説得力はありますか？

- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。
- プロットが説得力を持つかの目安は、全 CI を通る滑らかな線が引けるかです。



- 説得力はどうでしょうか？ おそらく、それほどでもないでしょう！

- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。
- プロットが説得力を持つかの目安は、全 CI を通る滑らかな線が引けるかです。

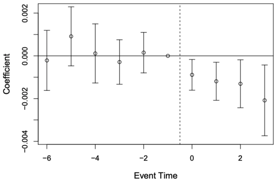


(a) Medicaid Eligibility

- こちらはどうですか？

- プレトレンドの点推定値だけでなく、信頼区間（CI）が何を排除できているかも重要です。
- プロットが説得力を持つかの目安は、全 CI を通る滑らかな線が引けるかです。

Figure 2: Effect of the ACA Medicaid Expansions on Annual Mortality



- そしてこちらは？

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が iid に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$$

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が *iid* に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i' \boldsymbol{\beta} + e_i$$

- 今は以下のような式を扱っています：

$$Y_{it} = \boldsymbol{X}_{it}' \boldsymbol{\beta} + e_{it}$$

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が iid に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$$

- 今は以下のような式を扱っています：

$$Y_{it} = \boldsymbol{X}_{it}'\boldsymbol{\beta} + e_{it}$$

- $(Y_{it}, \boldsymbol{X}_{it})$ が i と t を通じて iid であると仮定するのは妥当でしょうか？

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が iid に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i' \boldsymbol{\beta} + e_i$$

- 今は以下のような式を扱っています：

$$Y_{it} = \boldsymbol{X}_{it}' \boldsymbol{\beta} + e_{it}$$

- $(Y_{it}, \boldsymbol{X}_{it})$ が i と t を通じて iid であると仮定するのは妥当でしょうか？　いいえ。

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が iid に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$$

- 今は以下のような式を扱っています：

$$Y_{it} = \boldsymbol{X}_{it}'\boldsymbol{\beta} + e_{it}$$

- $(Y_{it}, \boldsymbol{X}_{it})$ が i と t を通じて iid であると仮定するのは妥当でしょうか？　いいえ。
  1. Y_{i1} と Y_{i2} は相関していることが予想されます。例えば、2010 年に高所得だった人は、2011 年も高所得である傾向があります。これは系列相関（serial autocorrelation）と呼ばれます。

パネル回帰の標準誤差

- $(Y_i, \boldsymbol{X}_i)$ が *iid* に抽出されている場合、以下の OLS 推定値の標準誤差を求める方法は既に知っています：

$$Y_i = \boldsymbol{X}_i'\boldsymbol{\beta} + e_i$$

- 今は以下のような式を扱っています：

$$Y_{it} = \boldsymbol{X}_{it}'\boldsymbol{\beta} + e_{it}$$

- $(Y_{it}, \boldsymbol{X}_{it})$ が i と t を通じて *iid* であると仮定するのは妥当でしょうか？　いいえ。
  1. Y_{i1} と Y_{i2} は相関していることが予想されます。例えば、2010 年に高所得だった人は、2011 年も高所得である傾向があります。これは系列相関（serial autocorrelation）と呼ばれます。
  2. さらに微妙な点として、処置が州レベルで割り当てられている場合、特定の州のすべての人は同じ D_{it} の値を持ちます（これは $\boldsymbol{X}_{it}$ に含まれます）。

クラスター標準誤差

- クラスター標準誤差は、同じ「クラスター」内の観測値間の相関を許容するように OLS 分散公式を拡張したものです。

クラスター標準誤差

- クラスター標準誤差は、同じ「クラスター」内の観測値間の相関を許容するように OLS 分散公式を拡張したものです。
- 仮定：各クラスターは独立にサンプリングされている。

クラスター標準誤差

- クラスター標準誤差は、同じ「クラスター」内の観測値間の相関を許容するように OLS 分散公式を拡張したものです。
- 仮定：各クラスターは独立にサンプリングされている。
- 例：個人レベル (i) でクラスター化する場合、Y_{i1} と Y_{i2} の依存は許容し、(Y_{i1}, Y_{i2}) と (Y_{j1}, Y_{j2}) は $j \neq i$ で独立。

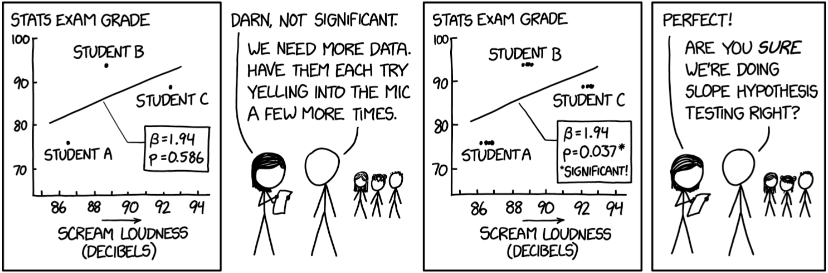
クラスター標準誤差

- クラスター標準誤差は、同じ「クラスター」内の観測値間の相関を許容するように OLS 分散公式を拡張したものです。
- 仮定：各クラスターは独立にサンプリングされている。
- 例：個人レベル (i) でクラスター化する場合、Y_{i1} と Y_{i2} の依存は許容し、(Y_{i1}, Y_{i2}) と (Y_{j1}, Y_{j2}) は $j \neq i$ で独立。
- パネル分析では最低でも個人レベルでクラスター化すべきです。処置が集計レベルで割り当てられているなら、そのレベルでクラスター化します。

クラスター標準誤差

- クラスター標準誤差は、同じ「クラスター」内の観測値間の相関を許容するように OLS 分散公式を拡張したものです。
- 仮定：各クラスターは独立にサンプリングされている。
- 例：個人レベル (i) でクラスター化する場合、Y_{i1} と Y_{i2} の依存は許容し、(Y_{i1}, Y_{i2}) と (Y_{j1}, Y_{j2}) は $j \neq i$ で独立。
- パネル分析では最低でも個人レベルでクラスター化すべきです。処置が集計レベルで割り当てられているなら、そのレベルでクラスター化します。
- 注意：中心極限定理に用いられる「実効的観測数」はクラスター数です。
    - クラスター数が非常に少ない（例：20 未満）場合は信頼できません。

XKCD


STATS EXAM GRADE
100
90
80
70
STUDENT B
STUDENT C
STUDENT A
β=1.94
p=0.586
86 88 90 92 94
SCREAM LOUDNESS (DECIBELS)
DARN, NOT SIGNIFICANT.
WE NEED MORE DATA. HAVE THEM EACH TRY YELLING INTO THE MIC A FEW MORE TIMES.
STATS EXAM GRADE
100
90
80
70
STUDENT B
STUDENT C
STUDENT A
β=1.94
p=0.037*
*SIGNIFICANT!
86 88 90 92 94
SCREAM LOUDNESS (DECIBELS)
PERFECT!
ARE YOU SURE WE'RE DOING SLOPE HYPOTHESIS TESTING RIGHT?

クラスター標準誤差の実装

- Stata でのクラスター標準誤差の実装は非常に簡単です。
- 以下を：
  reg y x, robust

  次のように置き換えるだけです：
  reg y x, cluster(clustervar)

```
. reg vehicle_fatality_rate beertax i.state i.year, r

Linear regression                               Number of obs     =        336
                                                F(54, 281)        =     128.89
                                                Prob > F          =     0.0000
                                                R-squared         =     0.9089
                                                Root MSE          =    1.9e-05

------------------------------------------------------------------------------
             |               Robust
vehicle_fa~e | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     beertax |   -.000064   .0000255    -2.51   0.013    -.0001141   -.0000139
             |
       state |
         AZ  |  -.0000547   .0000357    -1.53   0.127    -.0001249    .0000156
         AR  |  -.0000639   .0000293    -2.18   0.030    -.0001214   -6.26e-06
         CA  |  -.0001485   .0000412    -3.60   0.000    -.0002296   -.0000674
```

```
. reg vehicle_fatality_rate beertax i.state i.year, cluster(state)

Linear regression                               Number of obs     =        336
                                                F(6, 47)          =          .
                                                Prob > F          =          .
                                                R-squared         =     0.9089
                                                Root MSE          =    1.9e-05

                                 (Std. err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
vehicle_fa~e | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     beertax |   -.000064   .0000386    -1.66   0.104    -.0001416    .0000136
             |
       state |
         AZ  |  -.0000547   .0000506    -1.08   0.286    -.0001566    .0000472
         AR  |  -.0000639   .0000399    -1.60   0.116     -.000144    .0000163
         CA  |  -.0001485   .0000589    -2.52   0.015    -.0002671     -.00003
```

The Quarterly Journal of Economics, February 2004

HOW MUCH SHOULD WE TRUST DIFFERENCES-IN-DIFFERENCES ESTIMATES?*

MARIANNE BERTRAND
ESTHER DUFLO
SENDHIL MULLAINATHAN

Most papers that employ Differences-in-Differences estimation (DD) use many years of data and focus on serially correlated outcomes but ignore that the resulting standard errors are inconsistent. To illustrate the severity of this issue, we randomly generate placebo laws in state-level data on female wages from the Current Population Survey. For each law, we use OLS to compute the DD estimate of its "effect" as well as the standard error of this estimate. These conventional DD standard errors severely understate the standard deviation of the estimators: we find an "effect" significant at the 5 percent level for up to 45 percent of the placebo interventions. We use Monte Carlo simulations to investi-

非常に有名な DID

- Card and Krueger (1994) は、最低賃金が雇用にどのように影響するかを問いました。

非常に有名な DID

- Card and Krueger (1994) は、最低賃金が雇用にどのように影響するかを問いました。

- ミクロ経済学の理論で学んだことに基づけば、最低賃金の上昇は雇用にどう影響すると予想しますか？

非常に有名な DID

- Card and Krueger (1994) は、最低賃金が雇用にどのように影響するかを問いました。
- ミクロ経済学の理論で学んだことに基づけば、最低賃金の上昇は雇用にどう影響すると予想しますか？
  - 完全競争市場では、賃金（労働の価格）の下限設定は需要の減少を招くはずです。

非常に有名な DID

- Card and Krueger (1994) は、最低賃金が雇用にどのように影響するかを問いました。
- ミクロ経済学の理論で学んだことに基づけば、最低賃金の上昇は雇用にどう影響すると予想しますか？
  - 完全競争市場では、賃金（労働の価格）の下限設定は需要の減少を招くはずです。
- これを研究するため、CK は 1992 年にニュージャージー州が最低賃金を 4.25 ドルから 5.05 ドルに引き上げた事例を調査しました。

非常に有名な DID

- Card and Krueger (1994) は、最低賃金が雇用にどのように影響するかを問いました。
- ミクロ経済学の理論で学んだことに基づけば、最低賃金の上昇は雇用にどう影響すると予想しますか？
    - 完全競争市場では、賃金（労働の価格）の下限設定は需要の減少を招くはずです。
- これを研究するため、CK は 1992 年にニュージャージー州が最低賃金を 4.25 ドルから 5.05 ドルに引き上げた事例を調査しました。
- 彼らは、ニュージャージー州のファストフード店の雇用の変化を、最低賃金が 4.25 ドルのままだった隣接するペンシルベニア州の変化と比較する DID を用いました。

Variable		PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1.	FTE employment before,	23.33	20.44	-2.89
	all available observations	(1.35)	(0.51)	(1.44)
2.	FTE employment after,	21.17	21.03	-0.14
	all available observations	(0.94)	(0.52)	(1.07)
3.	Change in mean FTE	-2.16	0.59	2.76
	employment	(1.25)	(0.54)	(1.36)

Notes: Adapted from Card and Krueger (1994), Table 3. The table reports average full-time equivalent (FTE) employment at restaurants in Pennsylvania and New Jersey before and after a minimum wage increase in New Jersey. The sample consists of all stores with data on employment. Employment at six closed stores is set to zero. Employment at four temporarily closed stores is treated as missing. Standard errors are reported in parentheses

Variable	PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1. FTE employment before,	23.33	20.44	-2.89
all available observations	(1.35)	(0.51)	(1.44)
2. FTE employment after,	21.17	21.03	-0.14
all available observations	(0.94)	(0.52)	(1.07)
3. Change in mean FTE	-2.16	0.59	2.76
employment	(1.25)	(0.54)	(1.36)

Notes: Adapted from Card and Krueger (1994), Table 3. The table reports average full-time equivalent (FTE) employment at restaurants in Pennsylvania and New Jersey before and after a minimum wage increase in New Jersey. The sample consists of all stores with data on employment. Employment at six closed stores is set to zero. Employment at four temporarily closed stores is treated as missing. Standard errors are reported in parentheses

- 点推定値は 2.76 フルタイム当量（FTE）の雇用増加を示唆していますが、統計的に有意ではありませんでした。

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。
- この発見の一つの説明は、労働市場が完全競争的ではないというものです。むしろ、企業は雇い主独占（monopsonistic）的です。

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。
- この発見の一つの説明は、労働市場が完全競争的ではないというものです。むしろ、企業は雇い主独占（monopsonistic）的です。
  - 時給 7 ドルで 100 人の労働者を雇っている企業を考えます。
  - もう一人雇うと 10 ドルの追加利益が出るとします。しかし、そのためには賃金を時給 8 ドルに上げる必要があるとしましょう。

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。
- この発見の一つの説明は、労働市場が完全競争的ではないというものです。むしろ、企業は雇い主独占（monopsonistic）的です。
  - 時給 7 ドルで 100 人の労働者を雇っている企業を考えます。
  - もう一人雇うと 10 ドルの追加利益が出るとします。しかし、そのためには賃金を時給 8 ドルに上げる必要があるとしましょう。
  - 企業は賃金を 8 ドルに上げるべきでしょうか？

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。
- この発見の一つの説明は、労働市場が完全競争的ではないというものです。むしろ、企業は雇い主独占（monopsonistic）的です。
  - 時給 7 ドルで 100 人の労働者を雇っている企業を考えます。
  - もう一人雇うと 10 ドルの追加利益が出るとします。しかし、そのためには賃金を時給 8 ドルに上げる必要があるとしましょう。
  - 企業は賃金を 8 ドルに上げるべきでしょうか？ もし既存の 100 人全員に 1 ドル追加で払わなければならないなら、そうはしないでしょう！

なぜ？！

- 最低賃金の引き上げが雇用を減少させないように見えるという結果は、当時非常に驚き（かつ物議を醸すもの）でした。
- この発見の一つの説明は、労働市場が完全競争的ではないというものです。むしろ、企業は雇い主独占（monopsonistic）的です。
  - 時給 7 ドルで 100 人の労働者を雇っている企業を考えます。
  - もう一人雇うと 10 ドルの追加利益が出るとします。しかし、そのためには賃金を時給 8 ドルに上げる必要があるとしましょう。
  - 企業は賃金を 8 ドルに上げるべきでしょうか？ もし既存の 100 人全員に 1 ドル追加で払わなければならないなら、そうはしないでしょう！
  - しかし、もし最低賃金が 8 ドルに引き上げられれば、企業はどのみち最初の 100 人に 8 ドルを払わなければなりません。すると、追加の利益をもたらす 101 人目を 8 ドルで雇うことは理にかなうようになります。

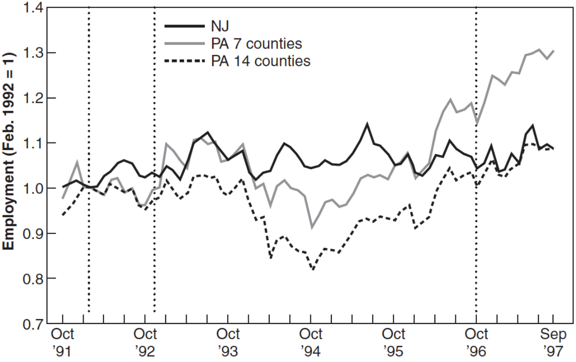


Figure 5.2.2 Employment in New Jersey and Pennsylvania fast food restaurants, October 1991 to September 1997 (from Card and Krueger 2000). Vertical lines indicate dates of the original Card and Krueger (1994) survey and the October 1996 federal minimum wage increase.

- 現代の基準で見れば、CK の分析はそれほど説得力があるとは言えないかもしれません。

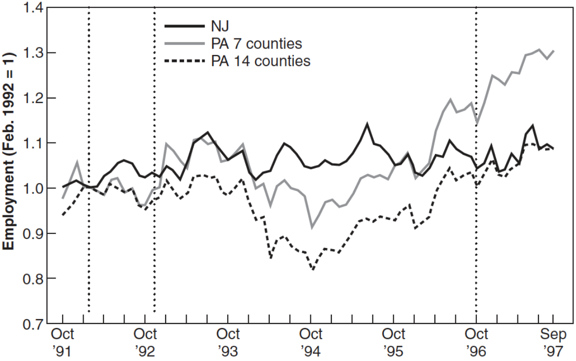


Figure 5.2.2 Employment in New Jersey and Pennsylvania fast food restaurants, October 1991 to September 1997 (from Card and Krueger 2000). Vertical lines indicate dates of the original Card and Krueger (1994) survey and the October 1996 federal minimum wage increase.

- 現代の基準で見れば、CK の分析はそれほど説得力があるとは言えないかもしれません。
- 1992 年 4 月の政策変更の前から、2 つの州は正確に並行して動いてはいません。

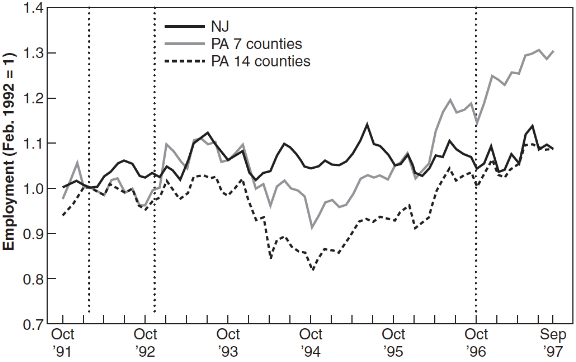


Figure 5.2.2 Employment in New Jersey and Pennsylvania fast food restaurants, October 1991 to September 1997 (from Card and Krueger 2000). Vertical lines indicate dates of the original Card and Krueger (1994) survey and the October 1996 federal minimum wage increase.

- 現代の基準で見れば、CK の分析はそれほど説得力があるとは言えないかもしれません。
- 1992 年 4 月の政策変更の前から、2 つの州は正確に並行して動いてはいません。また、対象が 2 つの州だけです！

処置タイミングのずれ

- 次に、最低賃金に関するより現代的な証拠を紹介します。
- その前に、処置のタイミングが分散している（staggered）設定、例えば州によって異なる年に最低賃金が導入される場合について DID を議論する必要があります。

処置タイミングのずれ

- 次に、最低賃金に関するより現代的な証拠を紹介します。
- その前に、処置のタイミングが分散している（staggered）設定、例えば州によって異なる年に最低賃金が導入される場合について DID を議論する必要があります。
- 約 5 年前までは、人々は以下のような OLS 回帰を実行することで、分散した設定に DID を拡張していました：

$$Y_{it} = \phi_i + \lambda_t + D_{it}\beta + e_{it}$$

ここで $D_{it} = 1$ は、ユニット i が期間 t で処置を受けていることを示します。

処置タイミングのずれ

- 次に、最低賃金に関するより現代的な証拠を紹介します。
- その前に、処置のタイミングが分散している（staggered）設定、例えば州によって異なる年に最低賃金が導入される場合について DID を議論する必要があります。
- 約 5 年前までは、人々は以下のような OLS 回帰を実行することで、分散した設定に DID を拡張していました：

$$Y_{it} = \phi_i + \lambda_t + D_{it}\beta + e_{it}$$

ここで $D_{it} = 1$ は、ユニット i が期間 t で処置を受けていることを示します。

- 2 期間モデルでは、これは処置群と対照群の標本平均の差の差に対応します。

処置タイミングのずれ

- 次に、最低賃金に関するより現代的な証拠を紹介します。
- その前に、処置のタイミングが分散している（staggered）設定、例えば州によって異なる年に最低賃金が導入される場合について DID を議論する必要があります。
- 約 5 年前までは、人々は以下のような OLS 回帰を実行することで、分散した設定に DID を拡張していました：

$$Y_{it} = \phi_i + \lambda_t + D_{it}\beta + e_{it}$$

  ここで $D_{it} = 1$ は、ユニット i が期間 t で処置を受けていることを示します。
- 2 期間モデルでは、これは処置群と対照群の標本平均の差の差に対応します。
- 残念ながら、タイミングが分散している場合、この推定量は処置群と未処置群の間の DID の平均にはならないことが判明しました。
  - Borusyak and Jaravel (2016), de Chaisemartin and D'Haultfoeuille (2020), Goodman-Bacon (2021) を参照してください。

- ここ数年、これらの回帰の問題を「修正」するための多くの研究が行われてきました。
- 解決策は通常、手作業で「クリーンな比較」を行うことです。

- ここ数年、これらの回帰の問題を「修正」するための多くの研究が行われてきました。
- 解決策は通常、手作業で「クリーンな比較」を行うことです。
  1. 年 g で初めて処置を受けたユニットについて、その期間中に処置を受けなかったユニットと比較して、$g-1$ と $g+k$ の間のアウトカムの変化を比較します。

- ここ数年、これらの回帰の問題を「修正」するための多くの研究が行われてきました。
- 解決策は通常、手作業で「クリーンな比較」を行うことです。
  1. 年 g で初めて処置を受けたユニットについて、その期間中に処置を受けなかったユニットと比較して、$g-1$ と $g+k$ の間のアウトカムの変化を比較します。
  2. これは、コホート g における処置 k 年後の効果の推定値となります。

- ここ数年、これらの回帰の問題を「修正」するための多くの研究が行われてきました。
- 解決策は通常、手作業で「クリーンな比較」を行うことです。
  1. 年 g で初めて処置を受けたユニットについて、その期間中に処置を受けなかったユニットと比較して、$g-1$ と $g+k$ の間のアウトカムの変化を比較します。
  2. これは、コホート g における処置 k 年後の効果の推定値となります。
  3. これをすべての g について行い、それらを合計して平均的な効果を求めます。

- ここ数年、これらの回帰の問題を「修正」するための多くの研究が行われてきました。

- 解決策は通常、手作業で「クリーンな比較」を行うことです。
  1. 年 g で初めて処置を受けたユニットについて、その期間中に処置を受けなかったユニットと比較して、$g-1$ と $g+k$ の間のアウトカムの変化を比較します。
  2. これは、コホート g における処置 k 年後の効果の推定値となります。
  3. これをすべての g について行い、それらを合計して平均的な効果を求めます。

- このアプローチや関連する手法には、Callaway and Sant'Anna (2020), Sun and Abraham (2020), Borusyak, Jaravel & Spiess (2021) など、多くの実装があります。

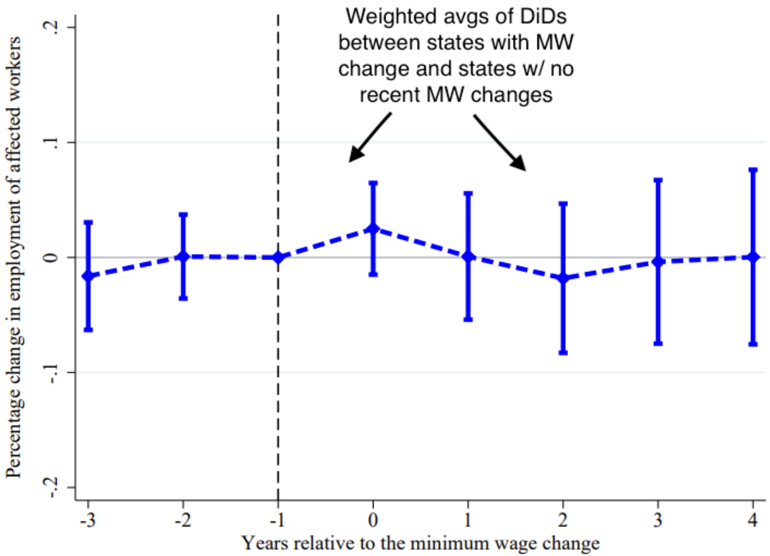
- Cengiz et al (2019) は、1976 年から 2016 年の間の 138 件の最低賃金変更を用いて、C&K の現代版を行いました。

- Cengiz et al (2019) は、1976 年から 2016 年の間の 138 件の最低賃金変更を用いて、C&K の現代版を行いました。
- 最低賃金を変更した各州について、前後 4 年間に最低賃金を変更しなかった州を「対照群」として抽出しました。

- Cengiz et al (2019) は、1976 年から 2016 年の間の 138 件の最低賃金変更を用いて、C&K の現代版を行いました。
- 最低賃金を変更した各州について、前後 4 年間に最低賃金を変更しなかった州を「対照群」として抽出しました。
- 彼らは、処置州と一致した対照州との間で DID を計算しました。

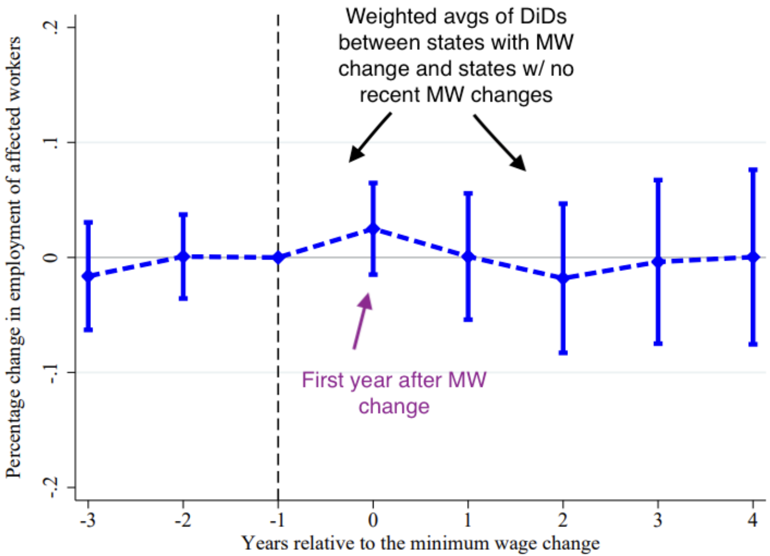
- Cengiz et al (2019) は、1976 年から 2016 年の間の 138 件の最低賃金変更を用いて、C&K の現代版を行いました。
- 最低賃金を変更した各州について、前後 4 年間に最低賃金を変更しなかった州を「対照群」として抽出しました。
- 彼らは、処置州と一致した対照州との間で DID を計算しました。
- そして、これらの DID の加重平均をとることで、全体的な平均効果を算出しました。

(a) Evolution of the missing and excess jobs



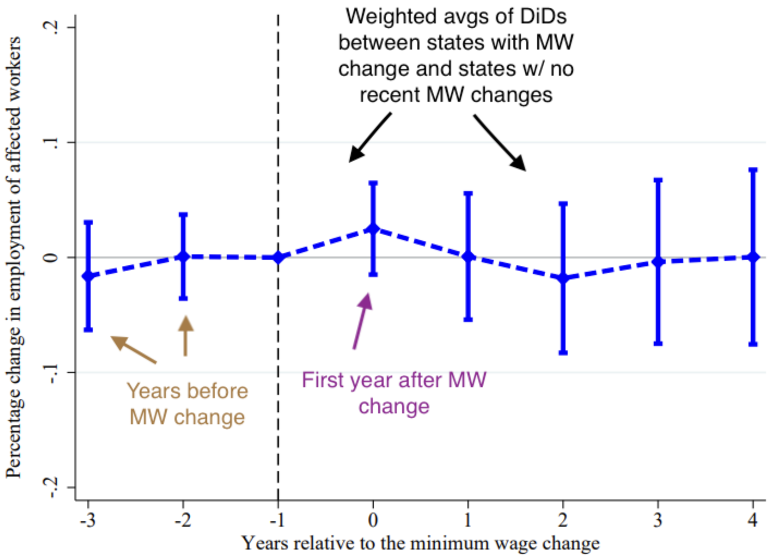
(b) Evolution of the employment of the affected workers

(a) Evolution of the missing and excess jobs



(b) Evolution of the employment of the affected workers

(a) Evolution of the missing and excess jobs



(b) Evolution of the employment of the affected workers

重要な考慮事項と注意点

- 歴史的な最低賃金の変更は、かなり緩やかなものでした。
  - 最低賃金を 4.25 ドルから 5.05 ドルに上げた変化が、7.25 ドルから 15 ドルへの引き上げについて示唆的であるかは不明です！

重要な考慮事項と注意点

- 歴史的な最低賃金の変更は、かなり緩やかなものでした。
  - 最低賃金を 4.25 ドルから 5.05 ドルに上げた変化が、7.25 ドルから 15 ドルへの引き上げについて示唆的であるかは不明です！
- 最低賃金の歴史的分析は、通常比較的短期的なものです。
  - 長期的なスパンでは、最低賃金の上昇は、労働者を代替するような技術へのシフトを促すかもしれません。

重要な考慮事項と注意点

- 歴史的な最低賃金の変更は、かなり緩やかなものでした。
  - 最低賃金を 4.25 ドルから 5.05 ドルに上げた変化が、7.25 ドルから 15 ドルへの引き上げについて示唆的であるかは不明です！
- 最低賃金の歴史的分析は、通常比較的短期的なものです。
  - 長期的なスパンでは、最低賃金の上昇は、労働者を代替するような技術へのシフトを促すかもしれません。
- 最低賃金が雇用を減少させるかどうかについては、経済学者の間でも依然として議論があります！

その他のパネルデータ手法

- ここでは、応用マイクロ経済学で最も一般的に使用されるパネルデータ手法である DID に焦点を当てました。
- しかし、他にも多くの手法があります：
  - ラグ付き依存変数のコントロール
  - 合成コントロール法 (Synthetic control)
  - 行列補完 (Matrix completion)
- これらをカバーする時間はありませんが、興味がある方は、さらに計量経済学の授業を履修することをお勧めします :)