Factorial Difference-in-Differences

UCLA CCPR

Yiqing Xu (Stanford)

Anqi Zhao (Duke)

Peng Ding (Berkeley)

Motivation

Motivation

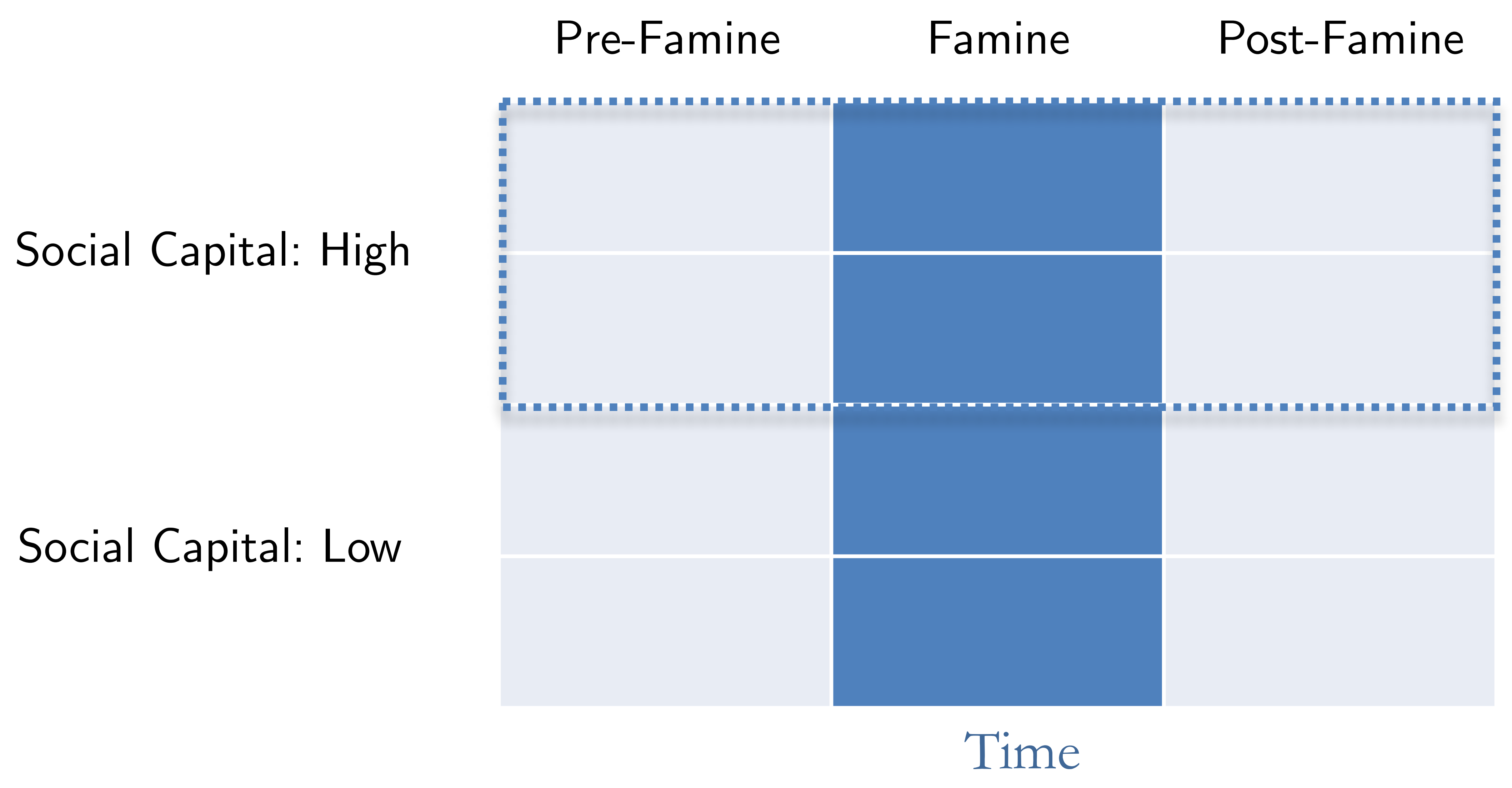
- Cao, Xu & Zhang (2022): “How social capital saved lives during China’s Great Famine”

Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure

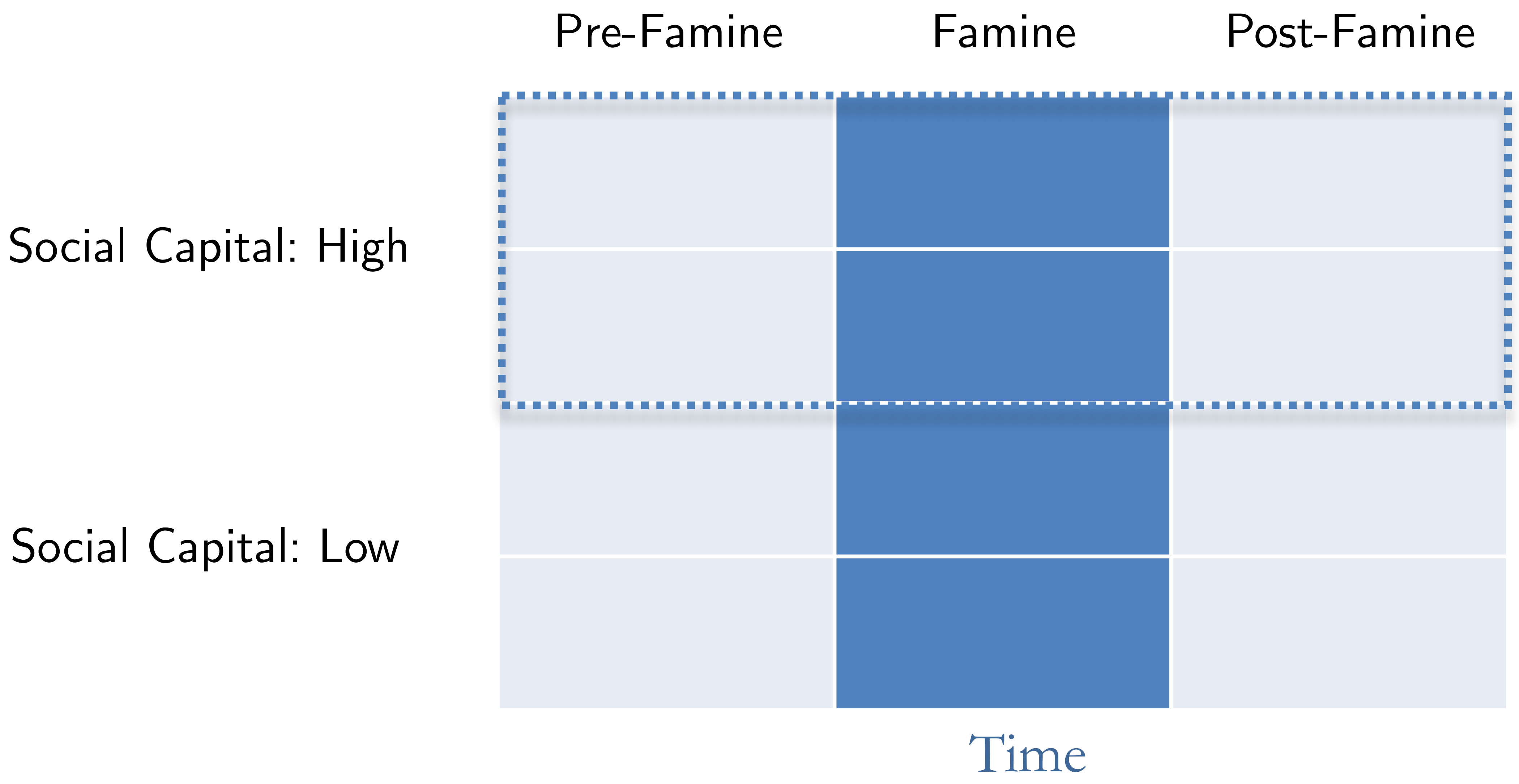
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure



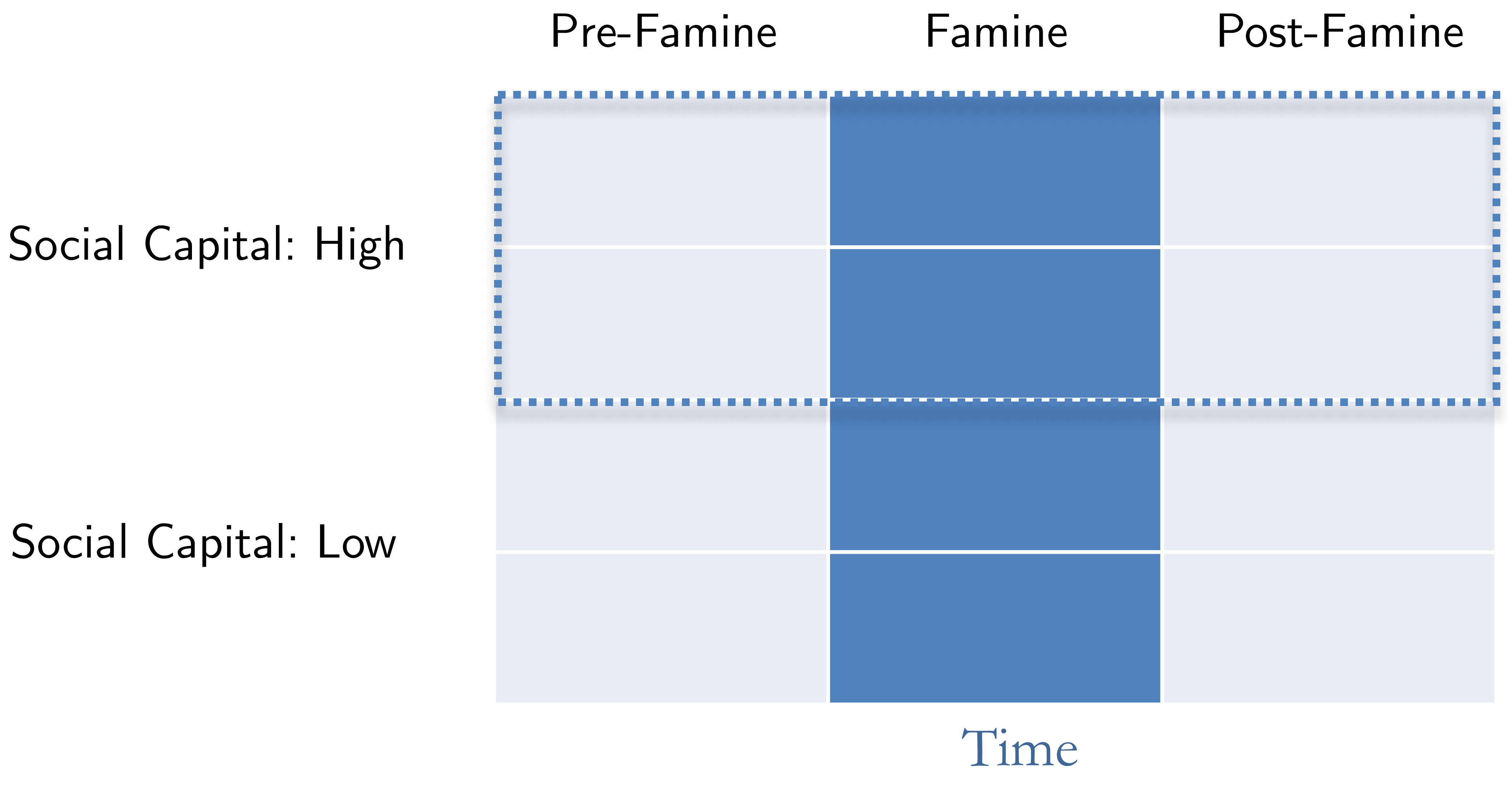
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital



Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}

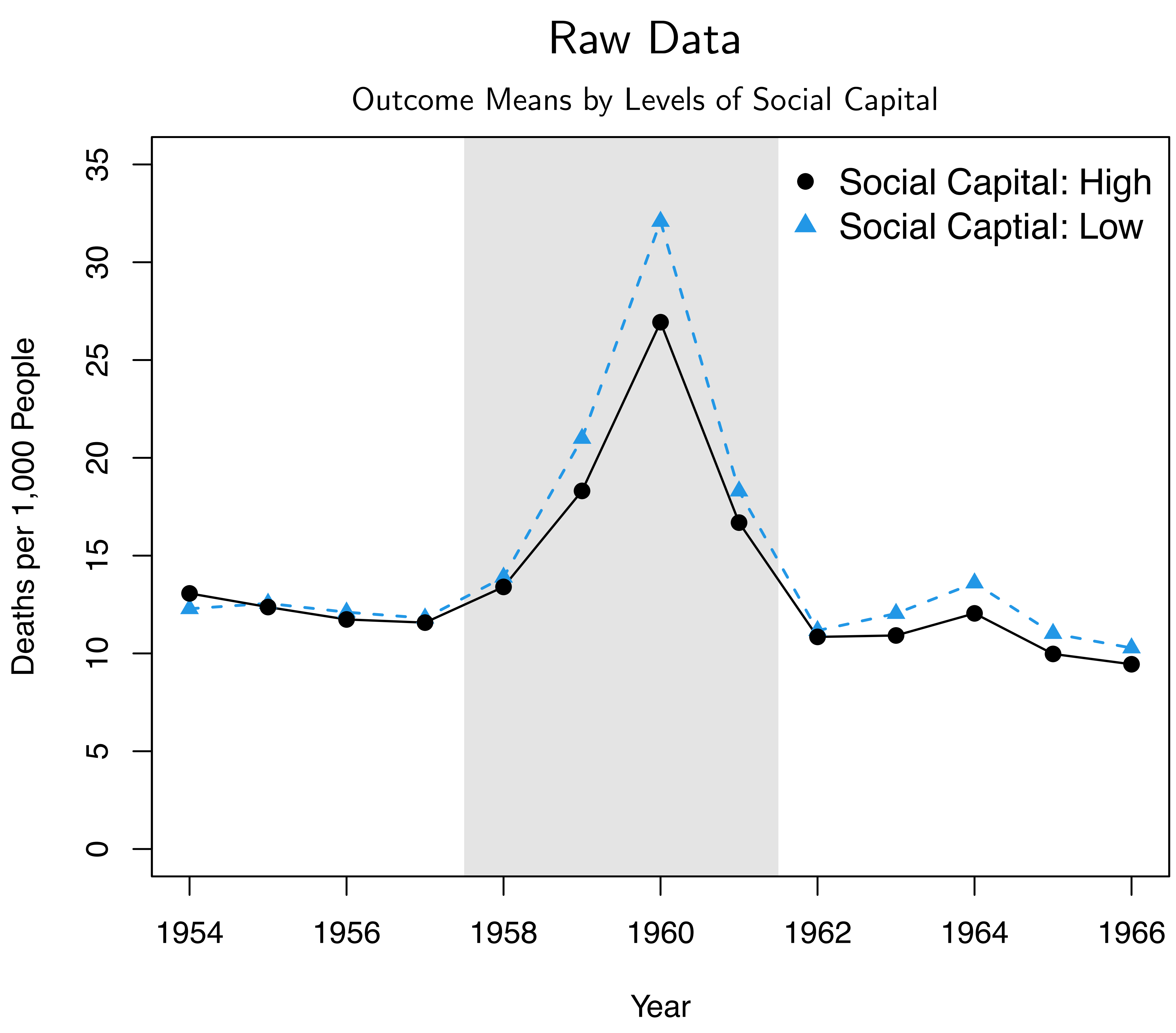


Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: $\{$Pre-Famine, Famine, Post-Famine$\}$

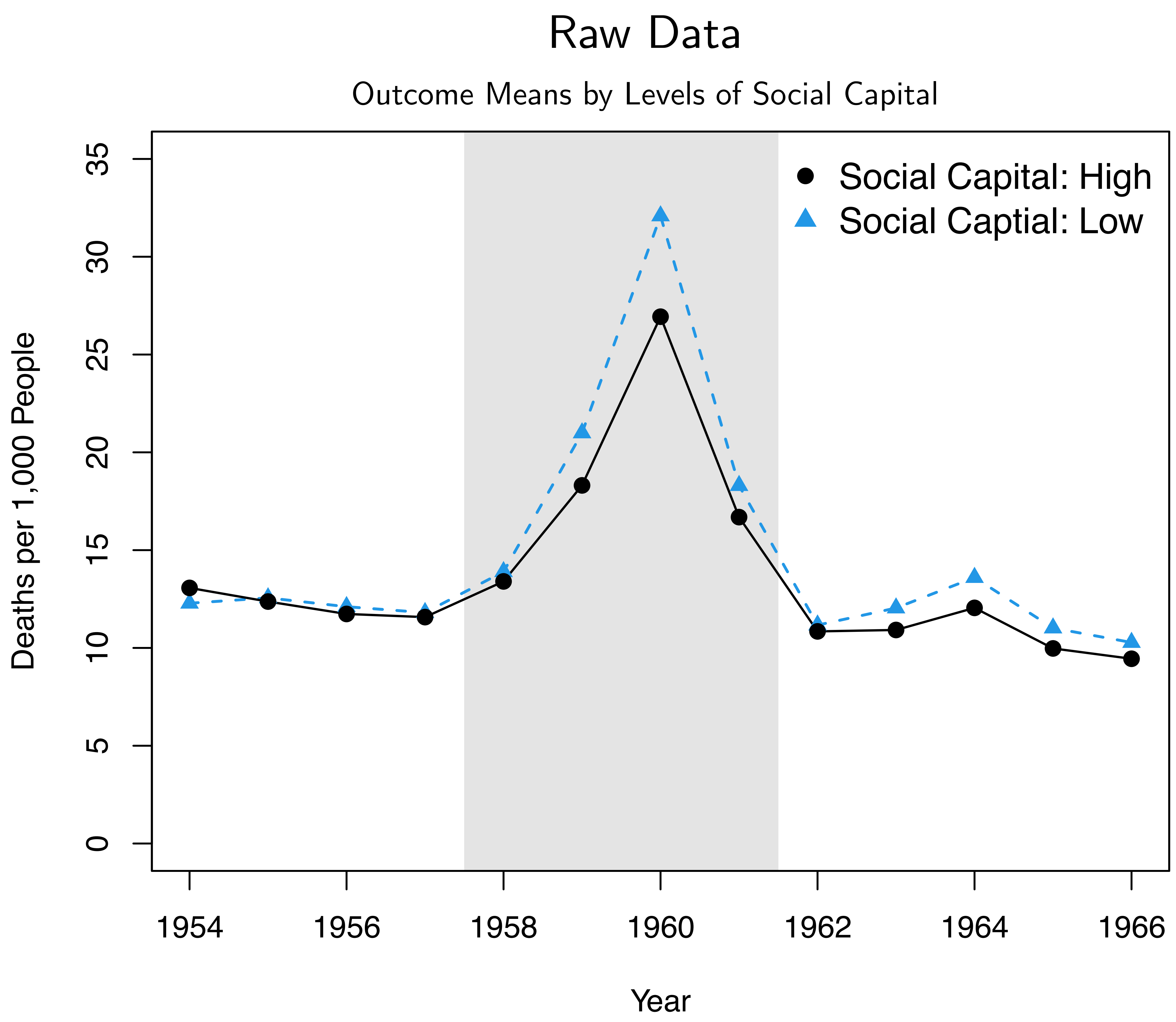
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}



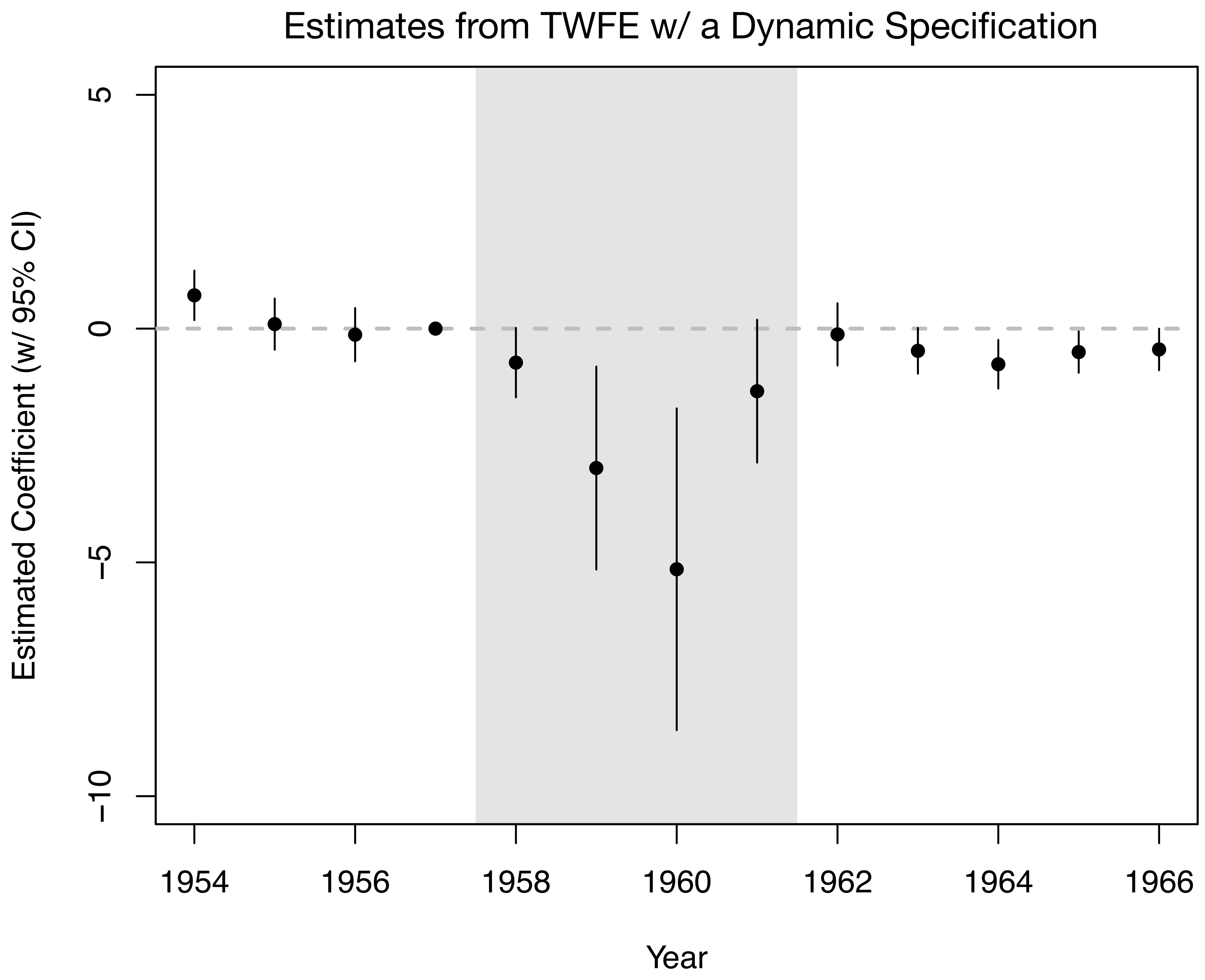
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}
- Estimation
  - Difference-in-differences (DID), or equivalently, two-way fixed effects (TWFE)



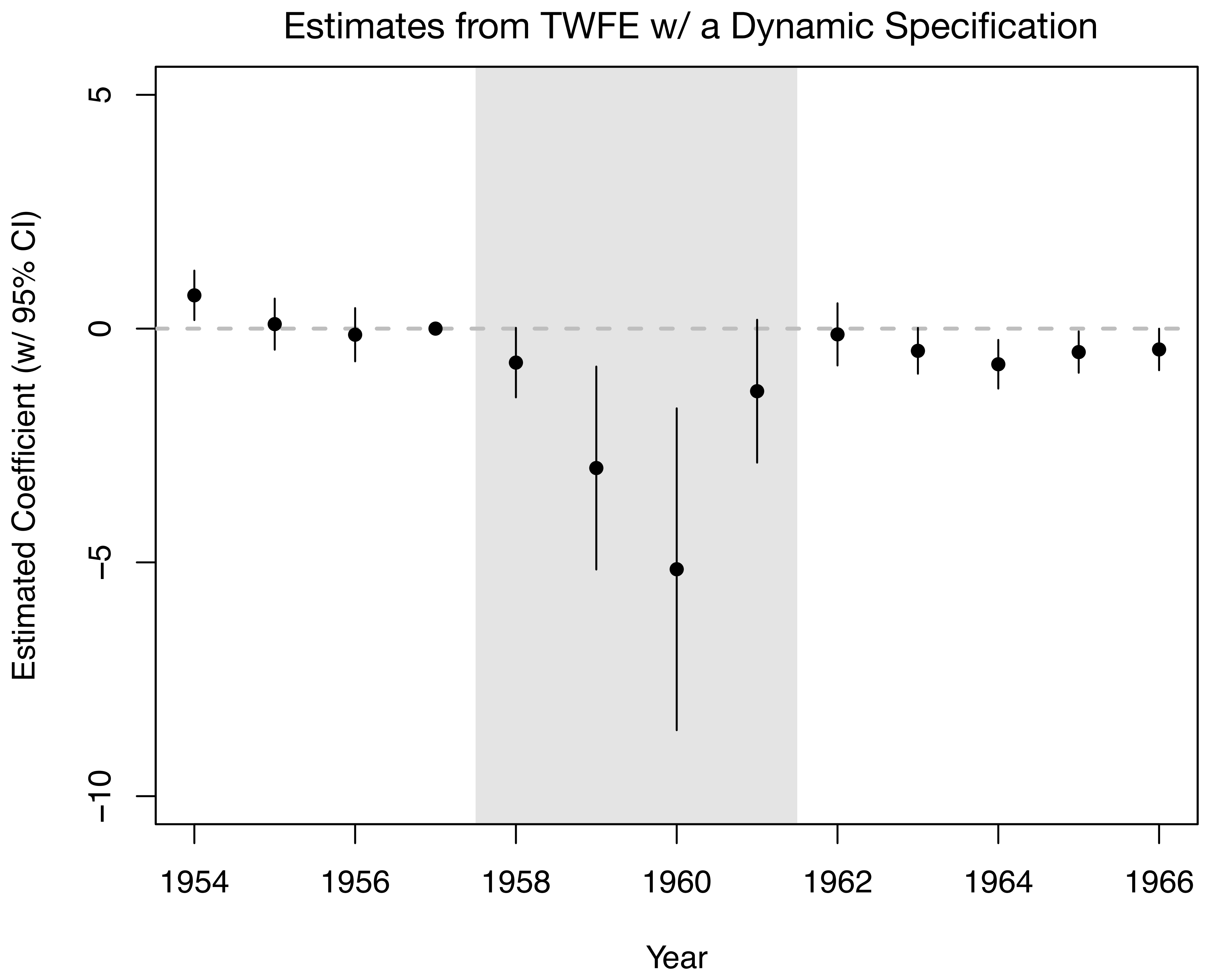
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}
- Estimation
  - Difference-in-differences (DID), or equivalently, two-way fixed effects (TWFE)



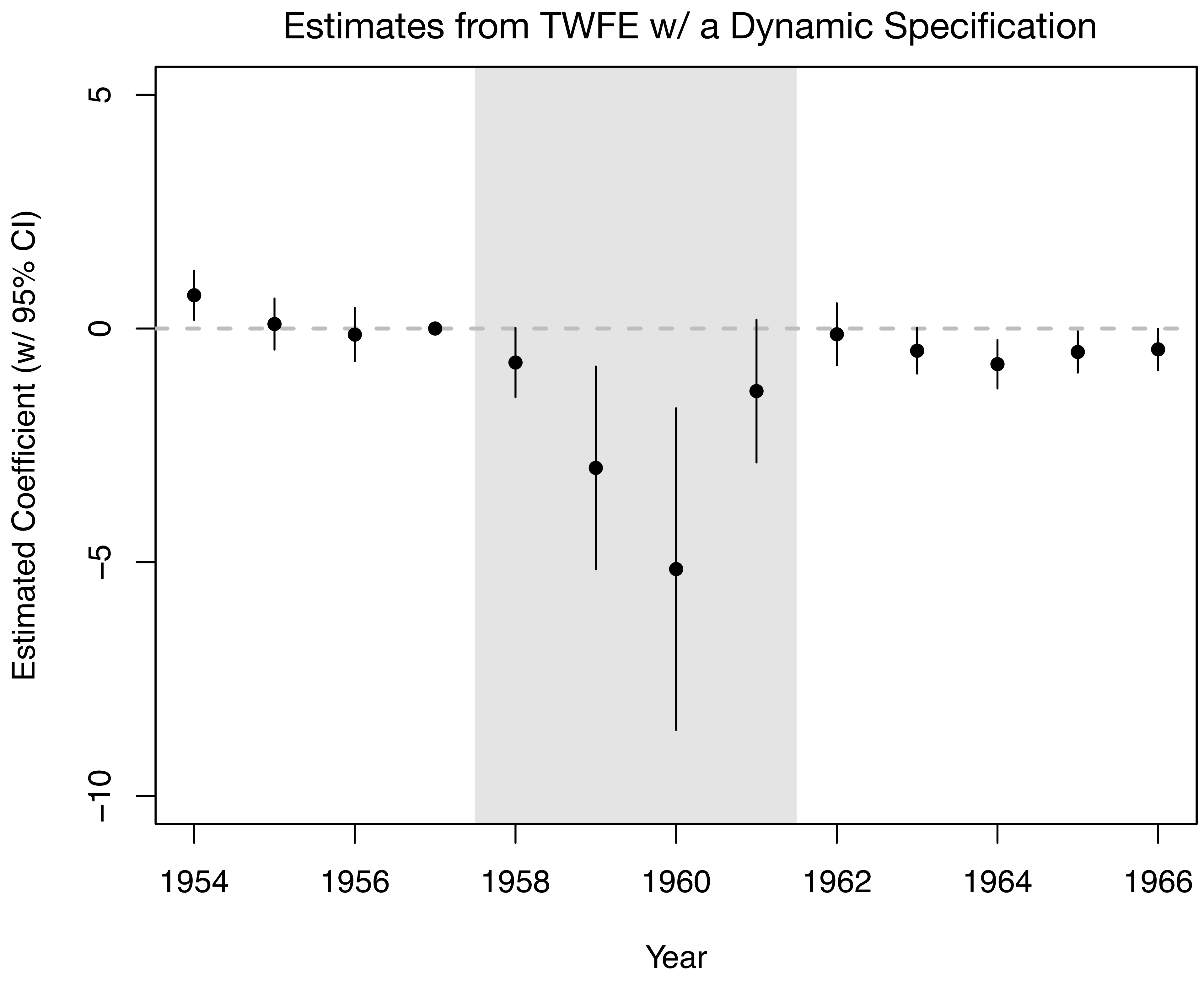
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}
- Estimation
  - Difference-in-differences (DID), or equivalently, two-way fixed effects (TWFE)
- Interpretation



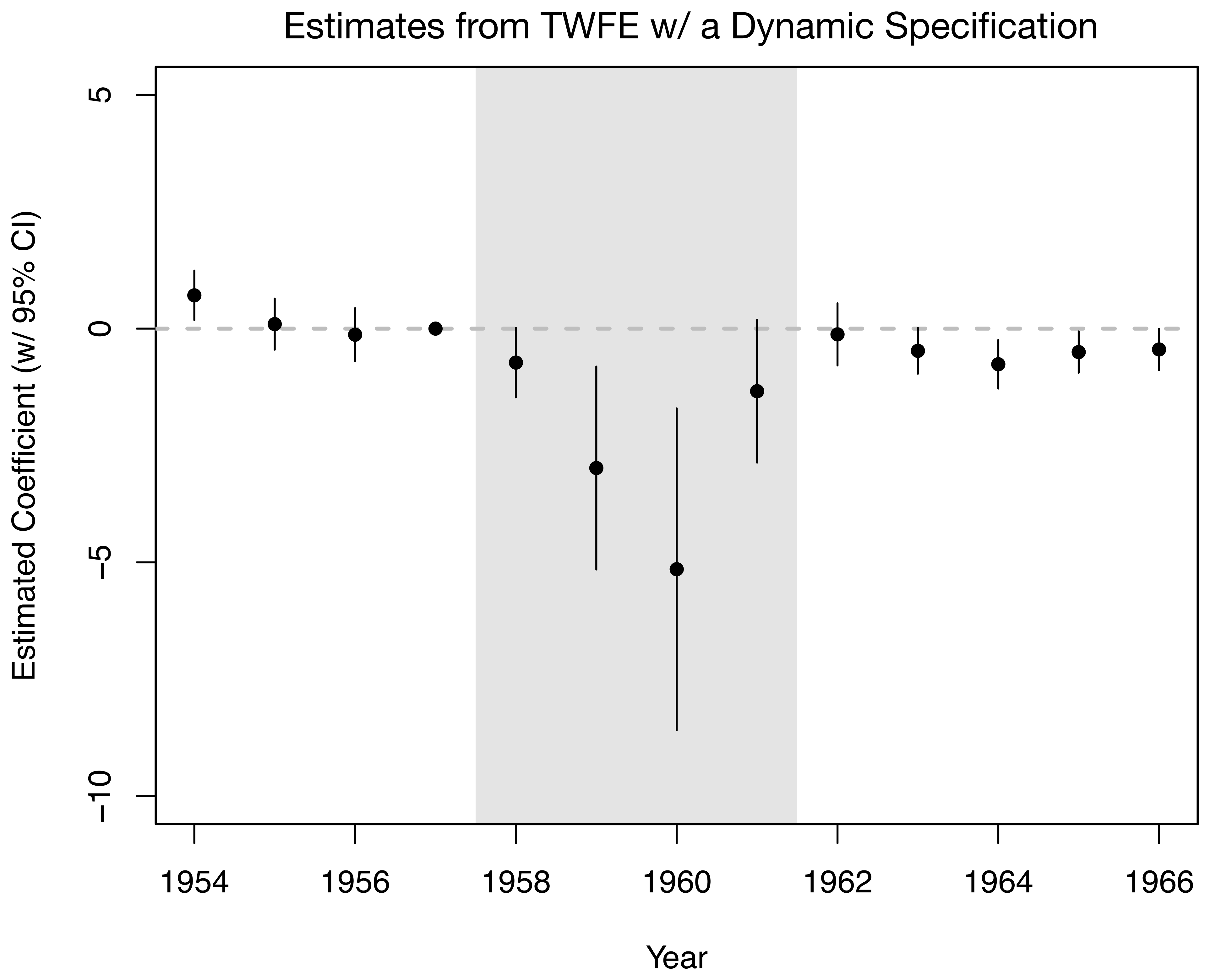
Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}
- Estimation
  - Difference-in-differences (DID), or equivalently, two-way fixed effects (TWFE)
- Interpretation
  - Descriptively: "the rise in the mortality rate during the famine years is significantly smaller in counties with higher social capital"



Motivation

- Cao, Xu & Zhang (2022): "How social capital saved lives during China's Great Famine"
- Data structure
  - A baseline factor G: time-invariant measure of social capital
  - Event time: {Pre-Famine, Famine, Post-Famine}
- Estimation
  - Difference-in-differences (DID), or equivalently, two-way fixed effects (TWFE)
- Interpretation
  - Descriptively: "the rise in the mortality rate during the famine years is significantly smaller in counties with higher social capital"
  - Causally: "we interpret these differences as the effects of social capital on famine relief."



More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA *Stanford University*

I study the effect of taste-based discrimination on the assimilation decisions of immigrant minorities. Do discriminated minority groups increase their assimilation efforts in order to avoid discrimination and public harassment or do they become alienated and retreat in their own communities? I exploit an exogenous shock to native attitudes, anti-Germanism in the United States during World War I, to empirically identify the reactions of German immigrants to increased native hostility. I use two measures of assimilation efforts: naming patterns and petitions for naturalization. In the face of increased discrimination, Germans increase their assimilation investments by Americanizing their own and their children's names and filing more petitions for US citizenship. These responses are stronger in states that registered higher levels of anti-German hostility, as measured by voting patterns and incidents of violence against Germans.

DOI: 10.1111/ajps.12888

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

From powerholders to stakeholders: State-building with elite compensation in early medieval China

Joy Chen[1] | Erik H. Wang[2] | Xiaoming Zhang[3]

[1]School of Economics, Renmin University of China, Beijing, China

[2]Wilf Family Department of Politics, New York University, New York, New York, USA

[3]Department of Sociology, Zhejiang University, Hangzhou, Zhejiang, China

Correspondence
Erik H. Wang, Wilf Family Department of Politics, New York University, New York, NY 10012, USA.
Email: erik.wang@nyu.edu

Authors contributed equally to this research and therefore share first-authorship equally. As far as "corresponding authorship" is concerned, each author is the corresponding author.

Abstract

How do rulers soften resistance by local powerholders to state-building efforts? This paper highlights a strategy of compensation, where elites receive government offices in exchange for relinquishing their localist interests, and become uprooted and integrated into the national political system as stakeholders. We explore this strategy in the context of the Northern Wei Dynasty of China (386–534 CE) that terminated an era of state weakness during which aristocrats exercised local autonomy through strongholds. Exploiting a comprehensive state-building reform in the late fifth century, we find that aristocrats from previously autonomous localities were disproportionately recruited into the bureaucracy as compensation for accepting stronger state presence. Three mechanisms of bureaucratic compensation facilitated state-building. Offices received by those aristocrats: (1) carried direct benefits, (2) realigned their interests toward the ruler, and (3) mitigated credible commitment problems. Our findings shed light on the "First Great Divergence" between Late Antiquity Europe and Medieval China.

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France†

By Mara P. Squicciarini*

This paper studies when religion can hamper diffusion of knowledge and economic development, and through which mechanism. I examine Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (*JEL* D83, I21, I26, N33, Z12)

EXPLAINING OUT-GROUP BIAS IN WEAK STATES
Religion and Legibility in the 1891/1892 Russian Famine

By VOLHA CHARNYSH

Department of Political Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
E-mail: charnysh@mit.edu

ABSTRACT

Two dominant explanations for ethnic bias in distributional outcomes are electoral incentives and out-group prejudice. This article proposes a novel and complementary explanation for the phenomenon: variation in legibility across ethnic groups. The author argues that states will allocate fewer resources to groups from which they cannot gather accurate information or collect taxes. The argument is supported by original data on state aid from the 1891/1892 famine in the Russian Empire. Qualitative and quantitative analyses show that districts with a larger Muslim population experienced higher famine mortality and received less generous public assistance. The Muslims, historically ruled via religious intermediaries, were less legible to state officials and generated lower fiscal revenues. State officials could not count on the repayment of food loans or collect tax arrears from Muslim communes, so they were more likely to withhold aid. State relief did not vary with the presence of other minorities that were more legible and generated more revenue.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA Stanford University

Discrimination against German immigrants × World War I

Do discriminated minority groups increase their assimilation efforts in order to avoid discrimination and public harassment or do they become alienated and retreat in their own communities? I exploit an exogenous shock to native attitudes, anti-Germanism in the United States during World War I, to empirically identify the reactions of German immigrants to increased native hostility. I use two measures of assimilation efforts: naming patterns and petitions for naturalization. In the face of increased discrimination, Germans increase their assimilation investments by Americanizing their own and their children's names and filing more petitions for US citizenship. These responses are stronger in states that registered higher levels of anti-German hostility, as measured by voting patterns and incidents of violence against Germans.

DOI: 10.1111/ajps.12888

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

From powerholders to stakeholders: State-building with elite compensation in early medieval China

Joy Chen[1] | **Erik H. Wang**[2] | **Xiaoming Zhang**[3]

[1]School of Economics, Renmin University of China, Beijing, China

[2]Wilf Family Department of Politics, New York University, New York, USA

[3]Department of Sociology, Zhejiang University, Hangzhou, Zhejiang, China

Correspondence
Erik H. Wang, Wilf Family Department of Politics, New York University, New York, NY 10012, USA.
Email: erik.wang@nyu.edu

Authors contributed equally to this research and therefore share first-authorship equally. As far as "corresponding authorship" is concerned, each author is the corresponding author.

Abstract
How do rulers soften resistance by local powerholders to state-building efforts? This paper highlights a strategy of compensation, where elites receive government offices in exchange for relinquishing their localist interests, and become uprooted and integrated into the national political system as stakeholders. We explore this strategy in the context of the Northern Wei Dynasty of China (386–534 CE) that terminated an era of state weakness during which aristocrats exercised local autonomy through strongholds. Exploiting a comprehensive state-building reform in the late fifth century, we find that aristocrats from previously autonomous localities were disproportionately recruited into the bureaucracy as compensation for accepting stronger state presence. Three mechanisms of bureaucratic compensation facilitated state-building. Offices received by those aristocrats: (1) carried direct benefits, (2) realigned their interests toward the ruler, and (3) mitigated credible commitment problems. Our findings shed light on the "First Great Divergence" between Late Antiquity Europe and Medieval China.

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France†

By MARA P. SQUICCIARINI*

This paper studies when religion can hamper diffusion of knowledge and economic development, and through which mechanism. I examine Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (*JEL* D83, I21, I26, N33, Z12)

EXPLAINING OUT-GROUP BIAS IN WEAK STATES
Religion and Legibility in the 1891/1892 Russian Famine

By VOLHA CHARNYSH

Department of Political Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
E-mail: charnysh@mit.edu

ABSTRACT
Two dominant explanations for ethnic bias in distributional outcomes are electoral incentives and out-group prejudice. This article proposes a novel and complementary explanation for the phenomenon: variation in legibility across ethnic groups. The author argues that states will allocate fewer resources to groups from which they cannot gather accurate information or collect taxes. The argument is supported by original data on state aid from the 1891/1892 famine in the Russian Empire. Qualitative and quantitative analyses show that districts with a larger Muslim population experienced higher famine mortality and received less generous public assistance. The Muslims, historically ruled via religious intermediaries, were less legible to state officials and generated lower fiscal revenues. State officials could not count on the repayment of food loans or collect tax arrears from Muslim communes, so they were more likely to withhold aid. State relief did not vary with the presence of other minorities that were more legible and generated more revenue.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA *Stanford University*

Discrimination against German immigrants × World War I

Do discriminated minority groups increase their assimilation efforts in order to avoid discrimination and public harassment or do they become alienated and retreat in their own communities? I exploit an exogenous shock to native attitudes, anti-Germanism in the United States during World War I, to empirically identify the reactions of German immigrants to increased native hostility. I use two measures of assimilation efforts: naming patterns and petitions for naturalization. In the face of increased discrimination, Germans increase their assimilation investments by Americanizing their own and their children's names and filing more petitions for US citizenship. These responses are stronger in states that registered higher levels of anti-German hostility, as measured by voting patterns and incidents of violence against Germans.

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France†

By Mara P. Squicciarini*

Catholicism × Industrial Revolution

This paper studies when religion can hamper diffusion of knowledge ... exam- the Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (*JEL* D83, I21, I26, N33, Z12)

DOI: 10.1111/ajps.12888

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

From powerholders to stakeholders: State-building with elite compensation in early medieval China

Joy Chen[1] | Erik H. Wang[2] | Xiaoming Zhang[3]

[1]School of Economics, Renmin University of China, Beijing, China
[2]Wilf Family Department of Politics, New York University, New York, New York, USA
[3]Department of Sociology, Zhejiang University, Hangzhou, Zhejiang, China

Correspondence
Erik H. Wang, Wilf Family Department of Politics, New York University, New York, NY 10012, USA.
Email: erik.wang@nyu.edu

Authors contributed equally to this research and therefore share first-authorship equally. As far as "corresponding authorship" is concerned, each author is the corresponding author.

Abstract
How do rulers soften resistance by local powerholders to state-building efforts? This paper highlights a strategy of compensation, where elites receive government offices in exchange for relinquishing their localist interests, and become uprooted and integrated into the national political system as stakeholders. We explore this strategy in the context of the Northern Wei Dynasty of China (386–534 CE) that terminated an era of state weakness during which aristocrats exercised local autonomy through strongholds. Exploiting a comprehensive state-building reform in the late fifth century, we find that aristocrats from previously autonomous localities were disproportionately recruited into the bureaucracy as compensation for accepting stronger state presence. Three mechanisms of bureaucratic compensation facilitated state-building. Offices received by those aristocrats: (1) carried direct benefits, (2) realigned their interests toward the ruler, and (3) mitigated credible commitment problems. Our findings shed light on the "First Great Divergence" between Late Antiquity Europe and Medieval China.

EXPLAINING OUT-GROUP BIAS IN WEAK STATES
Religion and Legibility in the 1891/1892 Russian Famine

By VOLHA CHARNYSH
Department of Political Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
E-mail: charnysh@mit.edu

ABSTRACT
Two dominant explanations for ethnic bias in distributional outcomes are electoral incentives and out-group prejudice. This article proposes a novel and complementary explanation for the phenomenon: variation in legibility across ethnic groups. The author argues that states will allocate fewer resources to groups from which they cannot gather accurate information or collect taxes. The argument is supported by original data on state aid from the 1891/1892 famine in the Russian Empire. Qualitative and quantitative analyses show that districts with a larger Muslim population experienced higher famine mortality and received less generous public assistance. The Muslims, historically ruled via religious intermediaries, were less legible to state officials and generated lower fiscal revenues. State officials could not count on the repayment of food loans or collect tax arrears from Muslim communes, so they were more likely to withhold aid. State relief did not vary with the presence of other minorities that were more legible and generated more revenue.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA Stanford University

Discrimination against German immigrants × World War I

... Do discriminated minority groups increase their assimilation efforts in order to avoid discrimination and public harassment or do they become alienated and retreat in their own communities? I exploit an exogenous shock to native attitudes, anti-Germanism in the United States during World War I, to empirically identify the reactions of German immigrants to increased native hostility. I use two measures of assimilation efforts: naming patterns and petitions for naturalization. In the face of increased discrimination, Germans increase their assimilation investments by Americanizing their own and their children's names and filing more petitions for US citizenship. These responses are stronger in states that registered higher levels of anti-German hostility, as measured by voting patterns and incidents of violence against Germans.

DOI: 10.1111/ajps.12888

ARTICLE

AMERICAN JOURNAL of POLITICAL SCIENCE

From powerholders to stakeholders: State-building with elite compensation in early medieval China

Joy Chen[1] | Erik H. Wang[2] | Xiaoming Zhang[3]

[1]School of Economics, Renmin University of China, Beijing, China
[2]Wilf Family Department of Politics, New York University, New York, NY
[3]Department of Sociology, Zhejiang University, Hangzhou, Zhejiang, China

Correspondence
Erik H. Wang, Wilf Family Department of Politics, New York University, New York, NY 10012, USA.
Email: erik.wang@nyu.edu

Authors contributed equally to this research and therefore share first-authorship equally. As far as "corresponding authorship" is concerned, each author is the corresponding author.

Elite Stronghold × State-building Reform

Abstract
How do rulers soften resistance by local powerholders to state-building ... receive government offices in exchange for relinquishing their localist interests, and become uprooted and integrated into the national political system as stakeholders. We explore this strategy in the context of the Northern Wei Dynasty of China (386–534 CE) that terminated an era of state weakness during which aristocrats exercised local autonomy through strongholds. Exploiting a comprehensive state-building reform in the late fifth century, we find that aristocrats from previously autonomous localities were disproportionately recruited into the bureaucracy as compensation for accepting stronger state presence. Three mechanisms of bureaucratic compensation facilitated state-building. Offices received by those aristocrats: (1) carried direct benefits, (2) realigned their interests toward the ruler, and (3) mitigated credible commitment problems. Our findings shed light on the "First Great Divergence" between Late Antiquity Europe and Medieval China.

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France[†]

By Mara P. Squicciarini*

Catholicism × Industrial Revolution

This paper studies when religion can hamper diffusion of knowledge ... exam- ... the Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (JEL D83, I21, I26, N33, Z12)

EXPLAINING OUT-GROUP BIAS IN WEAK STATES
Religion and Legibility in the 1891/1892 Russian Famine

By VOLHA CHARNYSH

Department of Political Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
E-mail: charnysh@mit.edu

ABSTRACT
Two dominant explanations for ethnic bias in distributional outcomes are electoral incentives and out-group prejudice. This article proposes a novel and complementary explanation for the phenomenon: variation in legibility across ethnic groups. The author argues that states will allocate fewer resources to groups from which they cannot gather accurate information or collect taxes. The argument is supported by original data on state aid from the 1891/1892 famine in the Russian Empire. Qualitative and quantitative analyses show that districts with a larger Muslim population experienced higher famine mortality and received less generous public assistance. The Muslims, historically ruled via religious intermediaries, were less legible to state officials and generated lower fiscal revenues. State officials could not count on the repayment of food loans or collect tax arrears from Muslim communes, so they were more likely to withhold aid. State relief did not vary with the presence of other minorities that were more legible and generated more revenue.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017 © American Political Science Association 2019

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA Stanford University

Discrimination against German immigrants × World War I

DOI: 10.1111/ajps.12888

ARTICLE

From powerholders to stakeholders: State-building with elite compensation in early medieval China

Joy Chen[1] | Erik H. Wang[2] | Xiaoming Zhang[3]

Elite Stronghold × State-building Reform

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France

By Mara P. Squicciarini

Catholicism × Industrial Revolution

EXPLAINING OUT-GROUP BIAS IN WEAK STATES

Religion and Legibility in the 1891/1892 Russian Famine

By VOLHA CHARNYSH

Muslim Share × 1891/1892 Russia Famine

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA Stanford University

I study the effect of taste-based discrimination on the assimilation decisions of immigrant minorities. Do discriminated minority groups increase their assimilation efforts in order to avoid discrimination and public harassment or do they become alienated and retreat in their own communities? I exploit an exogenous shock to native attitudes, anti-Germanism in the United States during World War I, to empirically identify the reaction of German immigrants to increased native hostility. I use two measures of assimilation efforts: naming patterns and petitions for naturalization. ... increase their assimilation investments by Americanizing their own and their children's names and by filing more petitions for US citizenship. These responses are stronger in states that registered higher levels of anti-German hostility, as measured by voting patterns and incidents of violence against Germans.

The Quarterly Journal of Economics (2011) 126, 593–650. doi:10.1093/qje/qjr009.
Advance Access publication on June 27, 2011.

THE POTATO'S CONTRIBUTION TO POPULATION AND URBANIZATION: EVIDENCE FROM A HISTORICAL EXPERIMENT*

NATHAN NUNN AND NANCY QIAN

We exploit regional variation in suitability for cultivating potatoes, together with time variation arising from their introduction to the Old World from the Americas, to estimate the impact of potatoes on Old World population and urbanization. Our results show that the introduction of the potato was responsible for a significant portion of the increase in population and urbanization observed during the eighteenth and nineteenth centuries. According to our most conservative estimates, the introduction of the potato accounts for approximately one-quarter of the growth in Old World population and urbanization ... and evidence from within-country comparisons of city populations and adult heights also confirms the cross-country findings. *JEL* Codes: J1, N1N5, O14.

Same DID Estimator, A Different Research Design

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France†

By Mara P. Squicciarini*

This paper studies when religion can hamper diffusion of knowledge and economic development, and through which mechanism. I examine Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (JEL D83, I21, I26, N33, Z12)

Econometrica, Vol. 84, No. 2 (March, 2016), 677–733

ELITE RECRUITMENT AND POLITICAL STABILITY: THE IMPACT OF THE ABOLITION OF CHINA'S CIVIL SERVICE EXAM

By Ying Bai and Ruixue Jia[1]

This paper studies how the abolition of an elite recruitment system—China's civil exam system that lasted over 1,300 years—affects political stability. Employing a panel data set across 262 prefectures and exploring the variations in the quotas on the entry-level exam candidates, we find that higher quotas per capita were associated with a higher probability of revolution participation after the abolition and a higher incidence of uprisings in 1911 that marked the end of the 2,000 years of imperial rule. This finding is robust to various checks including using the number of small rivers and short-run exam performance before the quota system as instruments. The patterns in the data appear most consistent with the interpretation that in regions with higher quotas per capita under the exam system, more would-be elites were negatively affected by the abolition. In addition, we document that modern human capital in the form of those studying in Japan also contributed to the revolution and that social capital strengthened the effect of quotas on revolution participation.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA *Stanford University*

I study the effect of taste-based discrimination on the assimilation decisions of immigrant minorities. Do discriminated minorit…

The Quarterly Journal of Economics (2011) 126, 593–650. doi:10.1093/qje/qjr009.
Advance Access publication on June 27, 2011.

THE POTATO'S CONTRIBUTION TO POPULATION AND URBANIZATION: EVIDENCE FROM A HISTORICAL EXPERIMENT*

NATHAN NUNN AND NANCY QIAN

We exploit regional variation in suitability for cultivating potatoes, together with time variation arising from their introduction to the Old World from the Americas, to estimate the impact of potatoes on Old World population and urbanization. Our results show that the introduction of the potato was responsible for a significant portion of the increase in population and urbaniza-…

Same DID Estimator, A Different Research Design

"Factorial Difference-in-Differences"

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France†

By Mara P. Squicciarini*

This paper studies when religion can hamper diffusion of knowledge and economic development, and through which mechanism. I examine Catholicism in France during the Second Industrial Revolution (1870–1914). In this period, technology became skill-intensive, leading to the introduction of technical education in primary schools. I find that more religious locations had lower economic development after 1870. Schooling appears to be the key mechanism: more religious areas saw a slower adoption of the technical curriculum and a push for religious education. In turn, religious education was negatively associated with industrial development 10 to 15 years later, when schoolchildren entered the labor market. (*JEL* D83, I21, I26, N33, Z12)

Econometrica, Vol. 84, No. 2 (March, 2016), 677–733

ELITE RECRUITMENT AND POLITICAL STABILITY: THE IMPACT OF THE ABOLITION OF CHINA'S CIVIL SERVICE EXAM

By Ying Bai and Ruixue Jia[1]

This paper studies how the abolition of an elite recruitment system—China's civil exam system that lasted over 1,300 years—affects political stability. Employing a panel data set across 262 prefectures and exploring the variations in the quotas on the entry-level exam candidates, we find that higher quotas per capita were associated with a higher probability of revolution participation after the abolition and a higher incidence of uprisings in 1911 that marked the end of the 2,000 years of imperial rule. This finding is robust to various checks including using the number of small rivers and short-run exam performance before the quota system as instruments. The patterns in the data appear most consistent with the interpretation that in regions with higher quotas per capita under the exam system, more would-be elites were negatively affected by the abolition. In addition, we document that modern human capital in the form of those studying in Japan also contributed to the revolution and that social capital strengthened the effect of quotas on revolution participation.

More Examples

American Political Science Review (2019) 113, 2, 405–422
doi:10.1017/S0003055419000017

How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I

VASILIKI FOUKA Stanford University

The Quarterly Journal of Economics (2011) 126, 593–650. doi:10.1093/qje/qjr009.
Advance Access publication on June 27, 2011.

THE POTATO'S CONTRIBUTION TO POPULATION AND URBANIZATION: EVIDENCE FROM A HISTORICAL EXPERIMENT*

NATHAN NUNN AND NANCY QIAN

Same DID Estimator, A Different Research Design

"Factorial Difference-in-Differences"

American Economic Review 2020, 110(11): 3454–3491
https://doi.org/10.1257/aer.20191054

Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France[†]

By MARA P. SQUICCIARINI*

Econometrica, Vol. 84, No. 2 (March, 2016), 677–733

ELITE RECRUITMENT AND POLITICAL STABILITY: THE IMPACT OF THE ABOLITION OF CHINA'S CIVIL SERVICE EXAM

BY YING BAI AND RUIXUE JIA[1]

This paper studies how the abolition of an elite recruitment system—China's civil exam system that lasted over 1,300 years—affects political stability. Employing a panel data set across 262 prefectures and exploring the variations in the quotas on the entry-level exam candidates, we find that higher quotas per capita were associated with a higher probability of revolution participation after the abolition and a higher incidence of uprisings in 1911 that marked the end of the 2,000 years of imperial rule. This finding is robust to various checks including using the number of small rivers and short-run exam performance before the quota system as instruments. The patterns in the data appear most consistent with the interpretation that in regions with higher quotas per capita under the exam system, more would-be elites were negatively affected by the abolition. In addition, we document that modern human capital in the form of those studying in Japan also contributed to the revolution and that social capital strengthened the effect of quotas on revolution participation.

- Factorial
  - classic topic in statistics
  - factorial experiments pioneered by R. A. Fisher and F. Yates
  - two treatment factors: their main effects and interaction are of interest

More Examples

Same DID Estimator, A Different Research Design

"Factorial Difference-in-Differences"

- Factorial
  - classic topic in statistics
  - factorial experiments pioneered by R. A. Fisher and F. Yates
  - two treatment factors: their main effects and interaction are of interest
- Difference-in-differences
  - popular in economics and related fields
  - a "research design" for causal inference with observational data
  - leverage panel data (units × times) to identify causal effects

This Paper

This Paper

- Factorial DID is a different research design from a canonical DID

This Paper

Setting, Estimand, Estimator,
ID Assumptions & ID Results

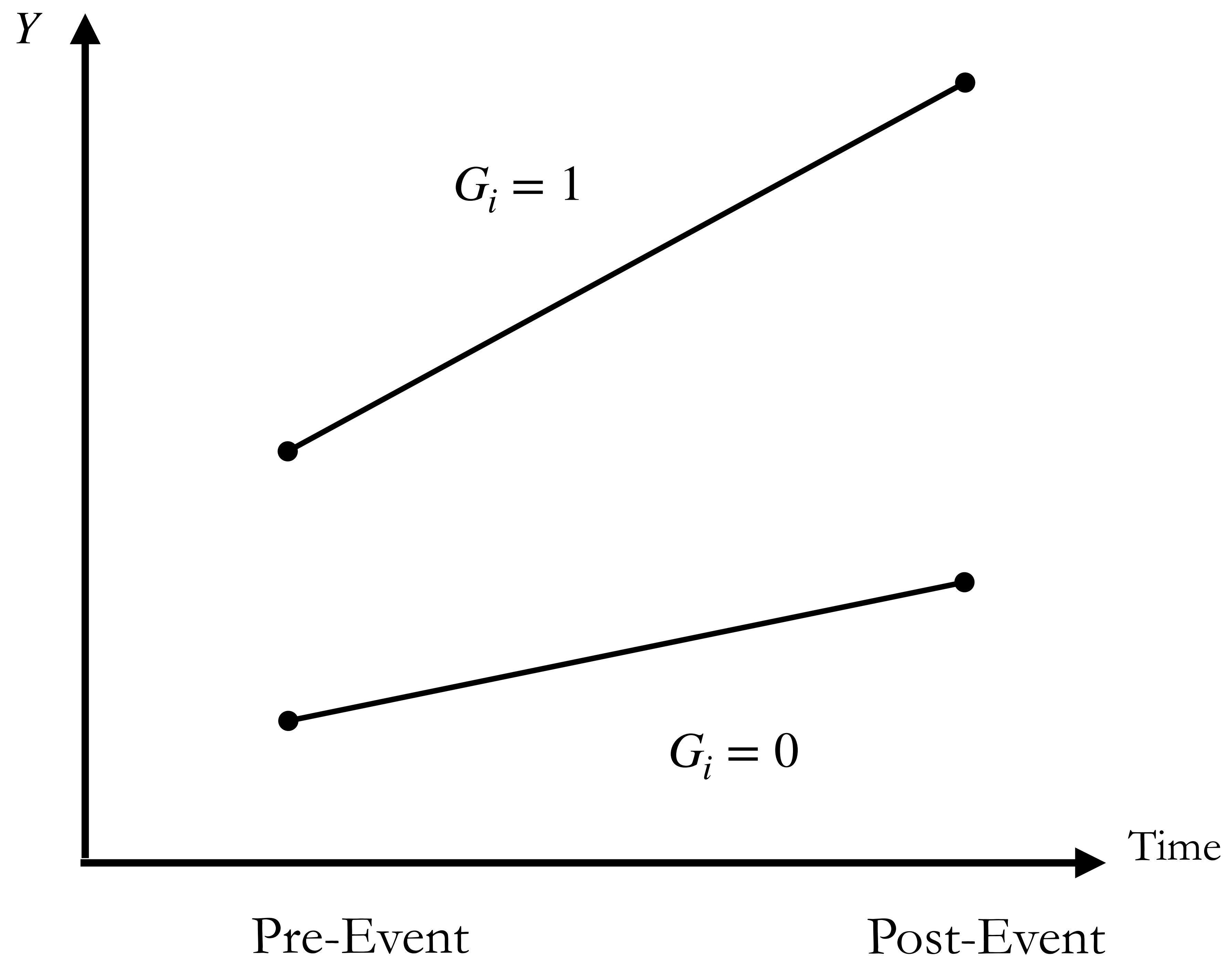
- Factorial DID is a different research design from a canonical DID

This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

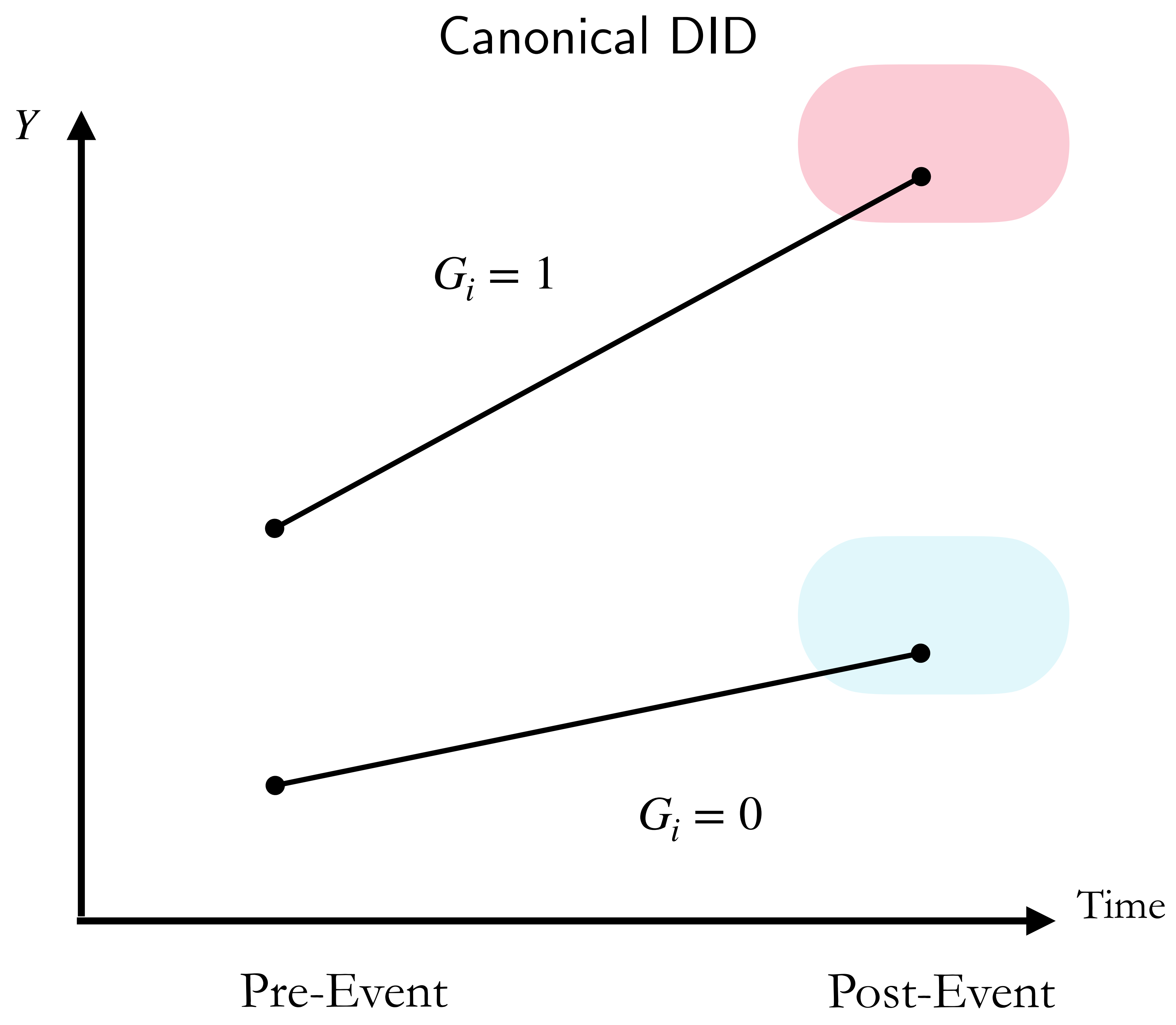
Canonical DID



This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

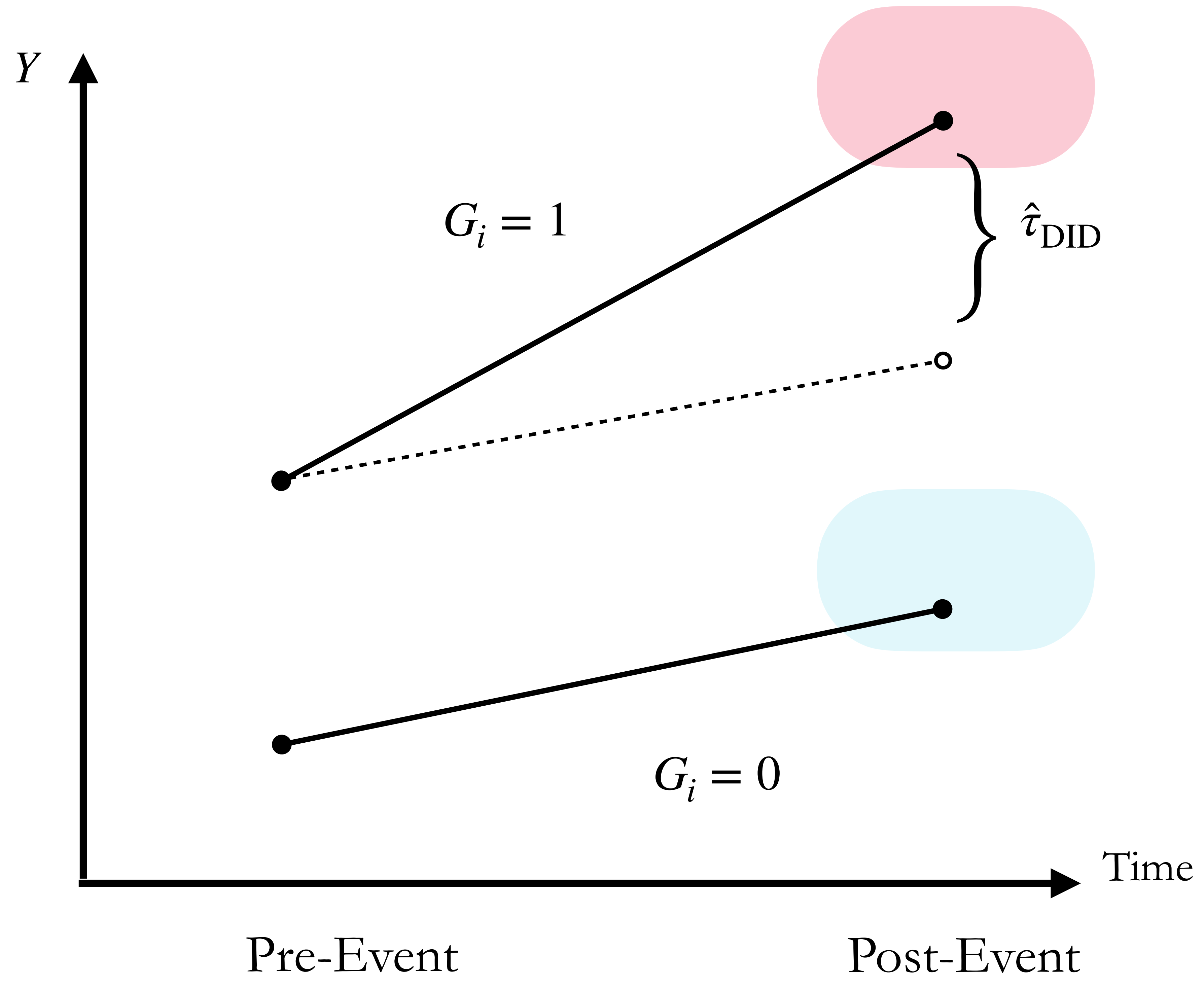


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

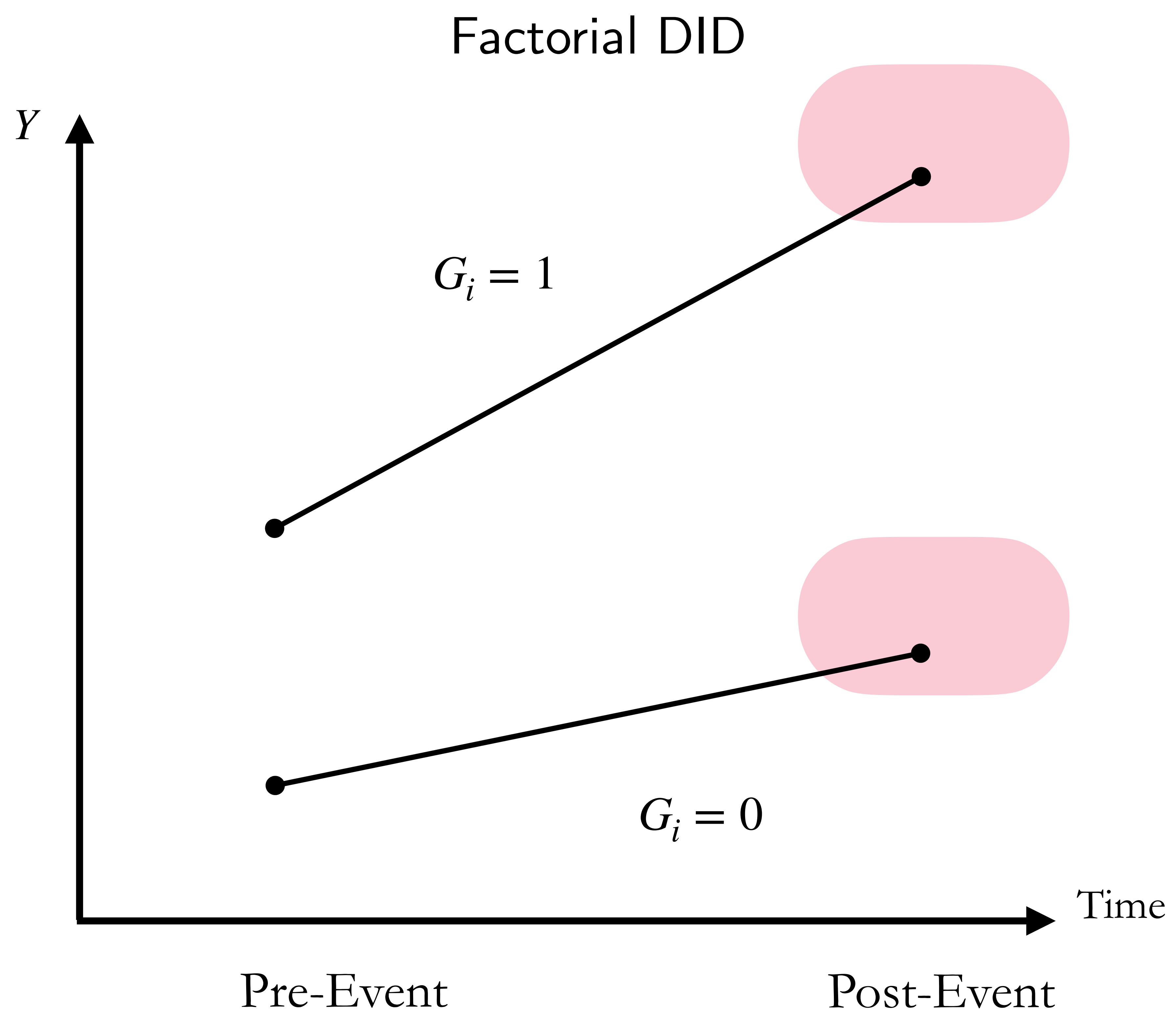
Canonical DID



This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

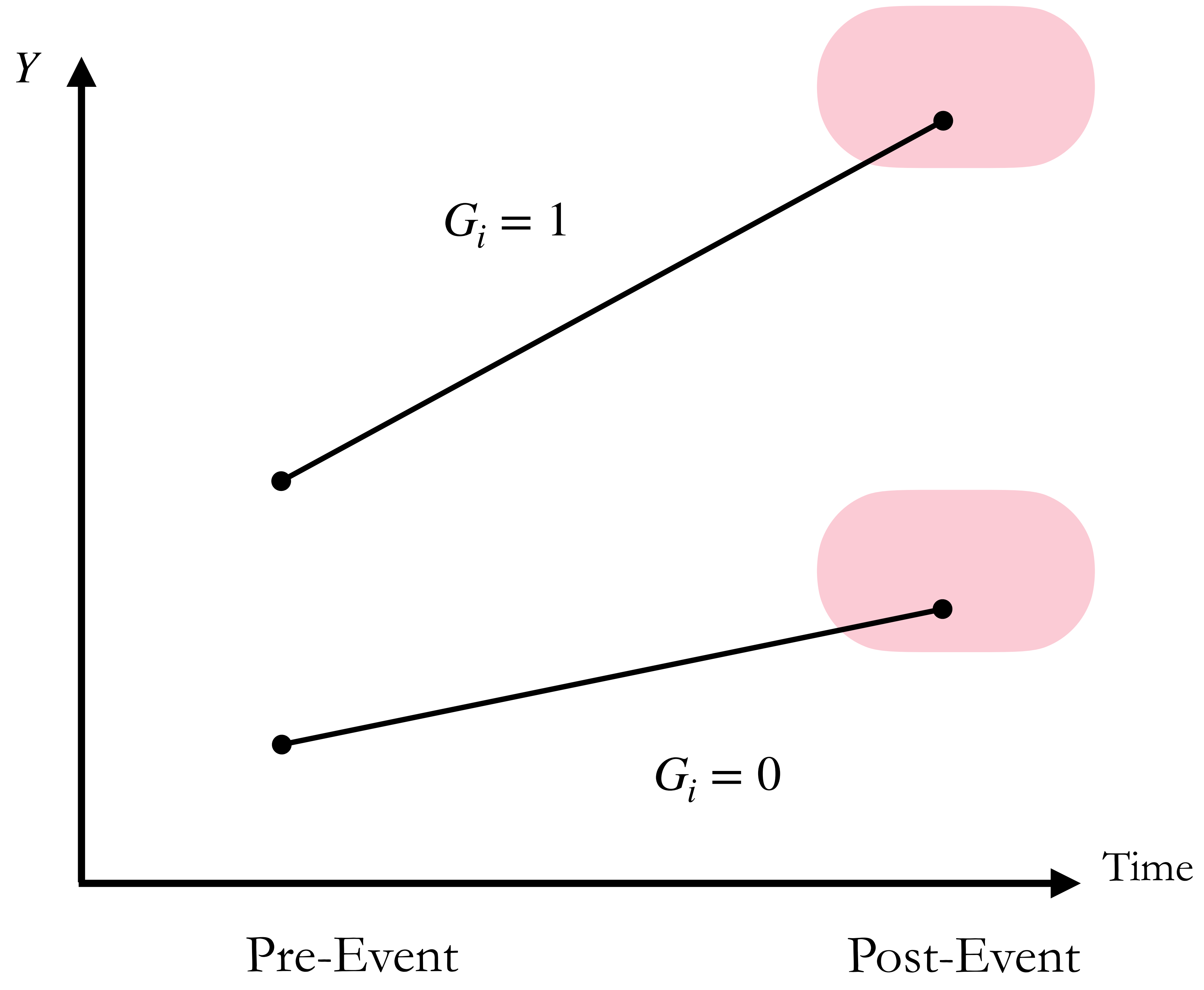


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event

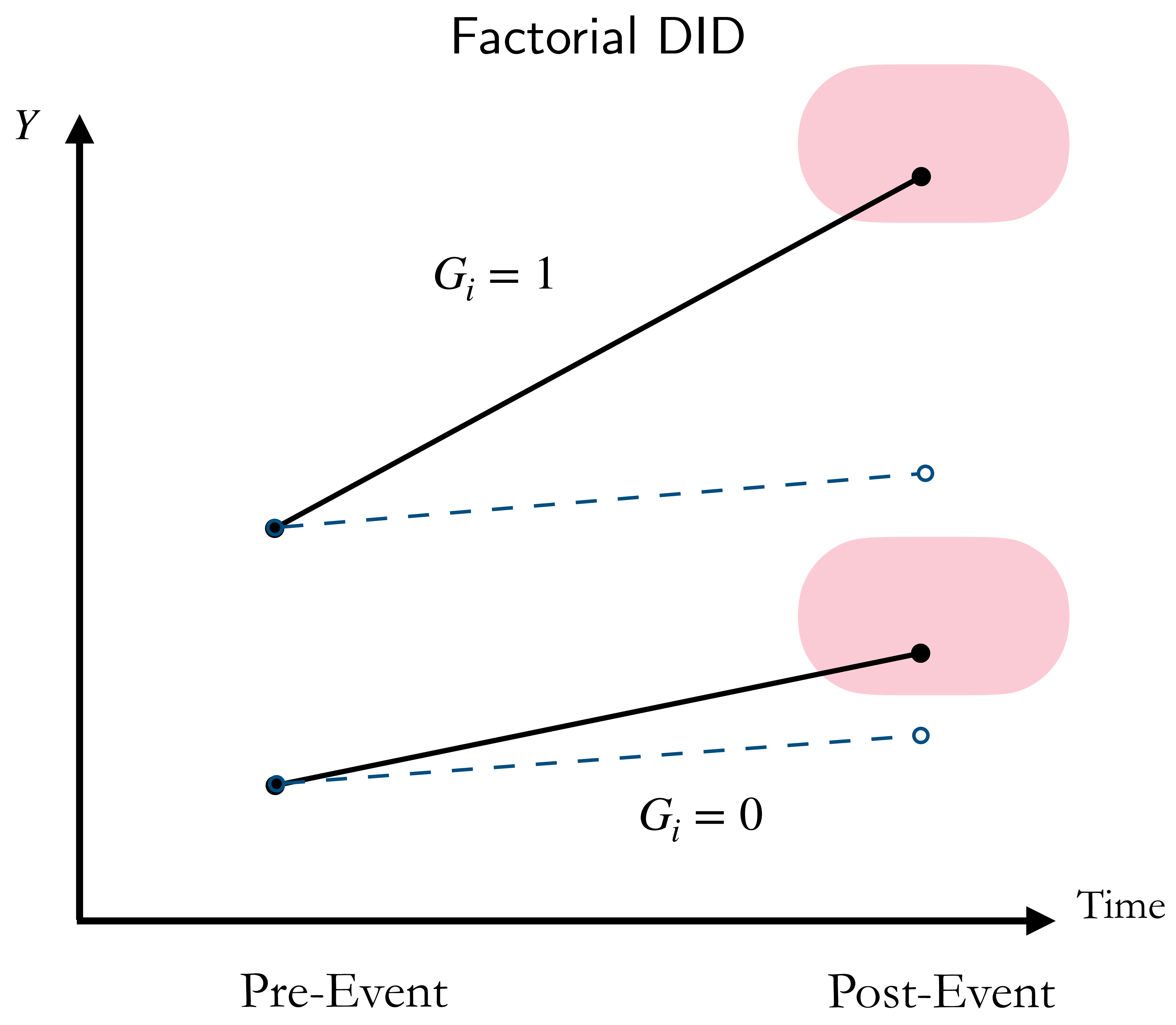


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event

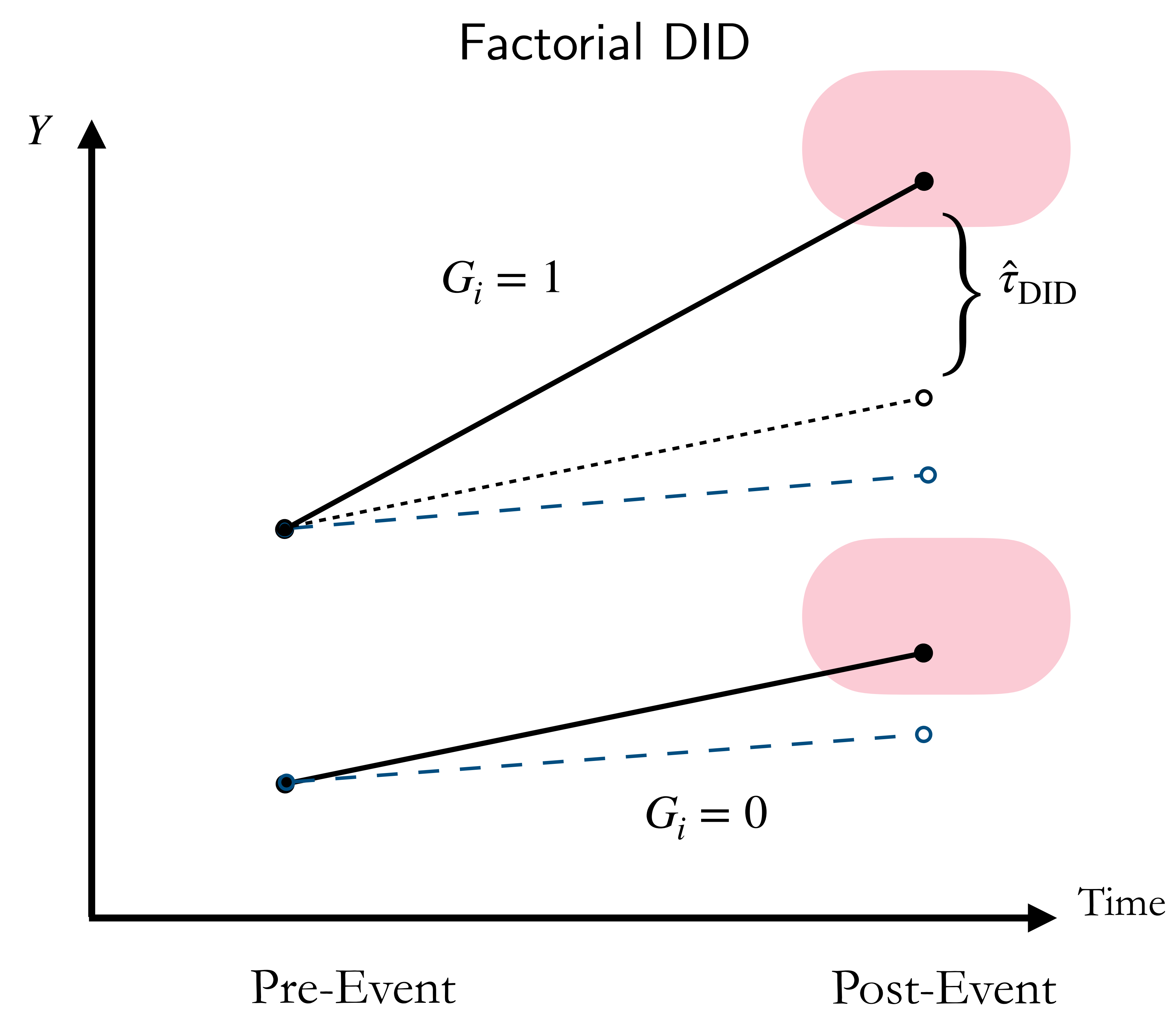


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event

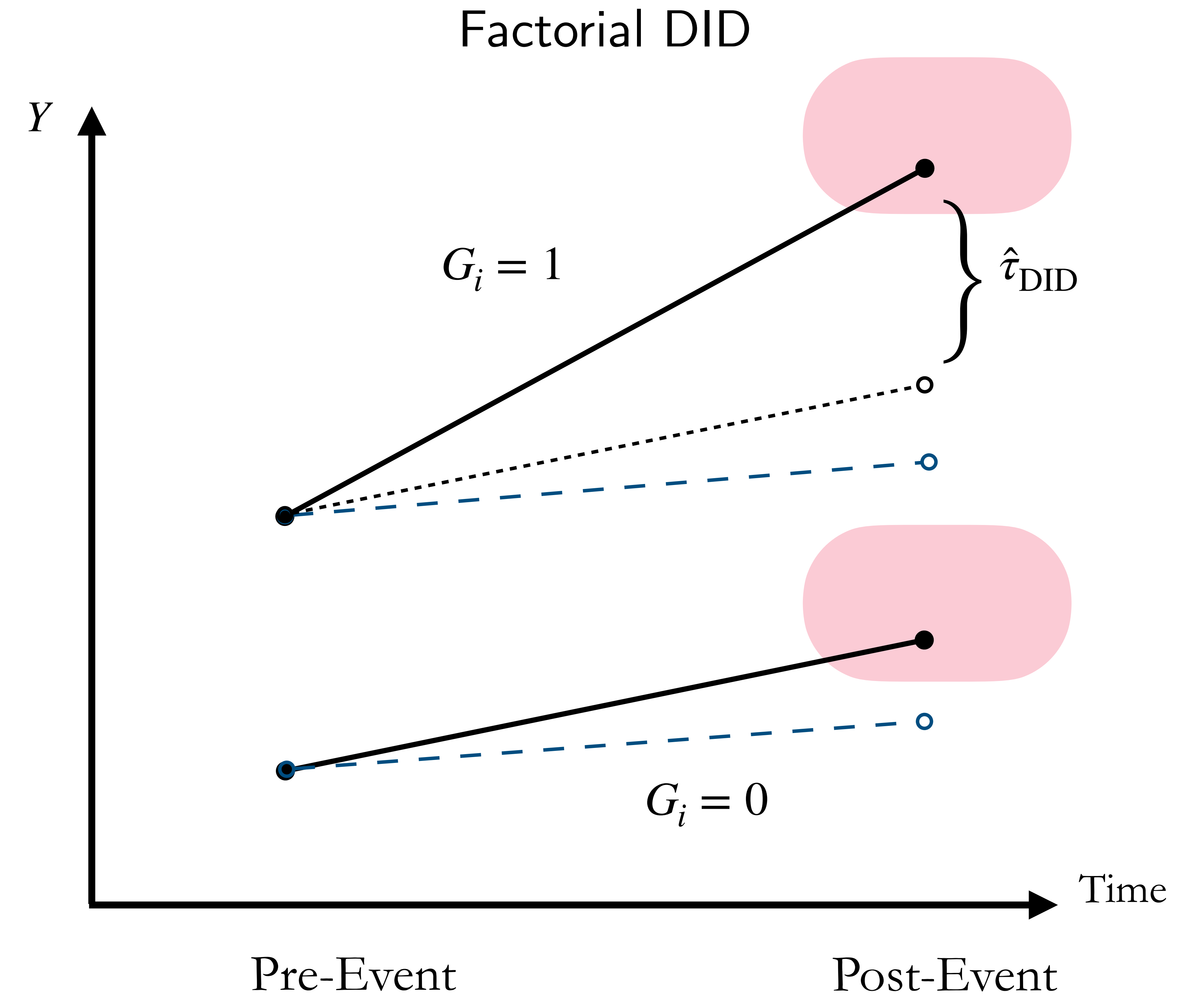


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event
  - Not causal

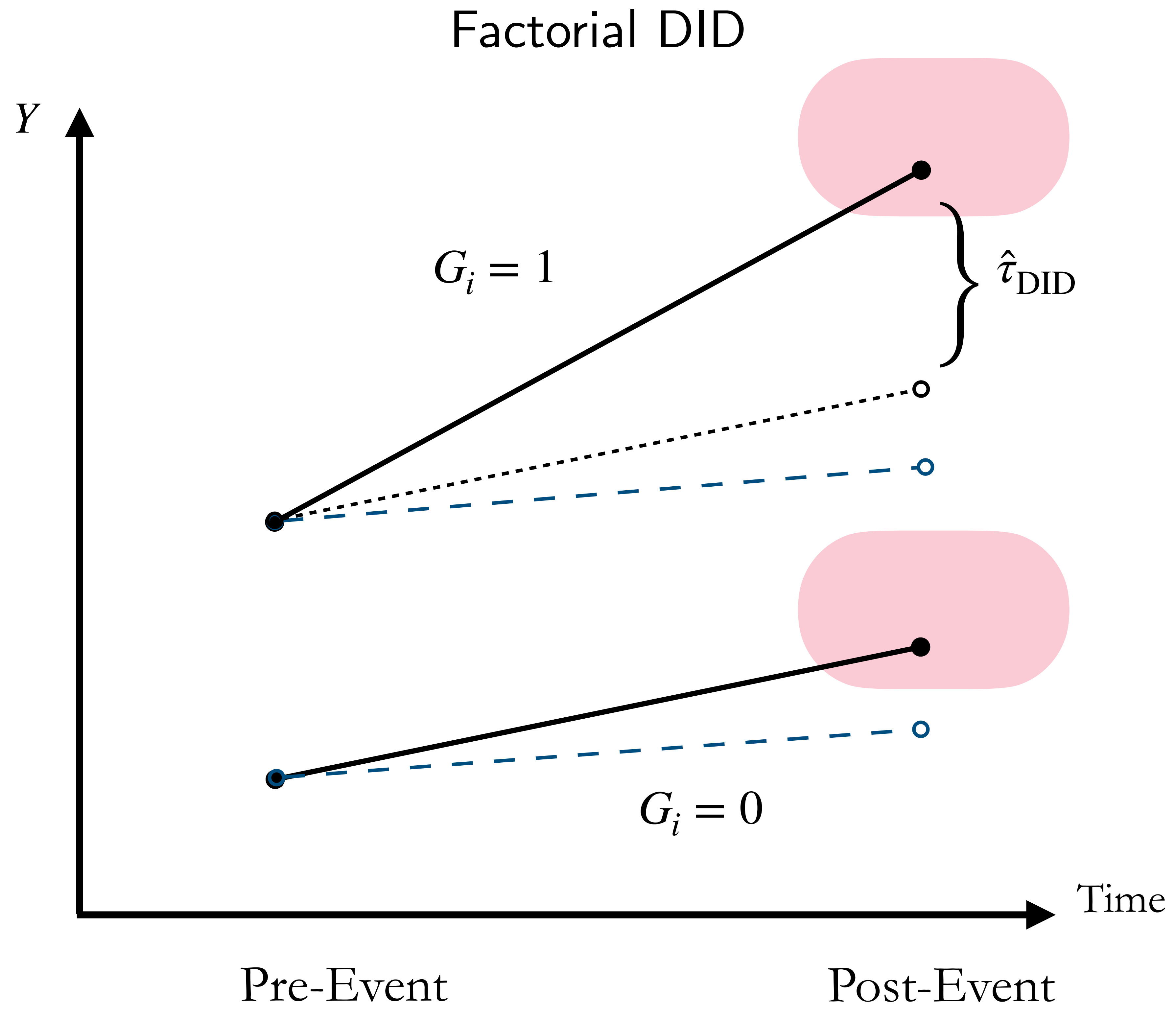


This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event
  - Not causal
  - Need additional analytical tools (factorial designs) to clarify

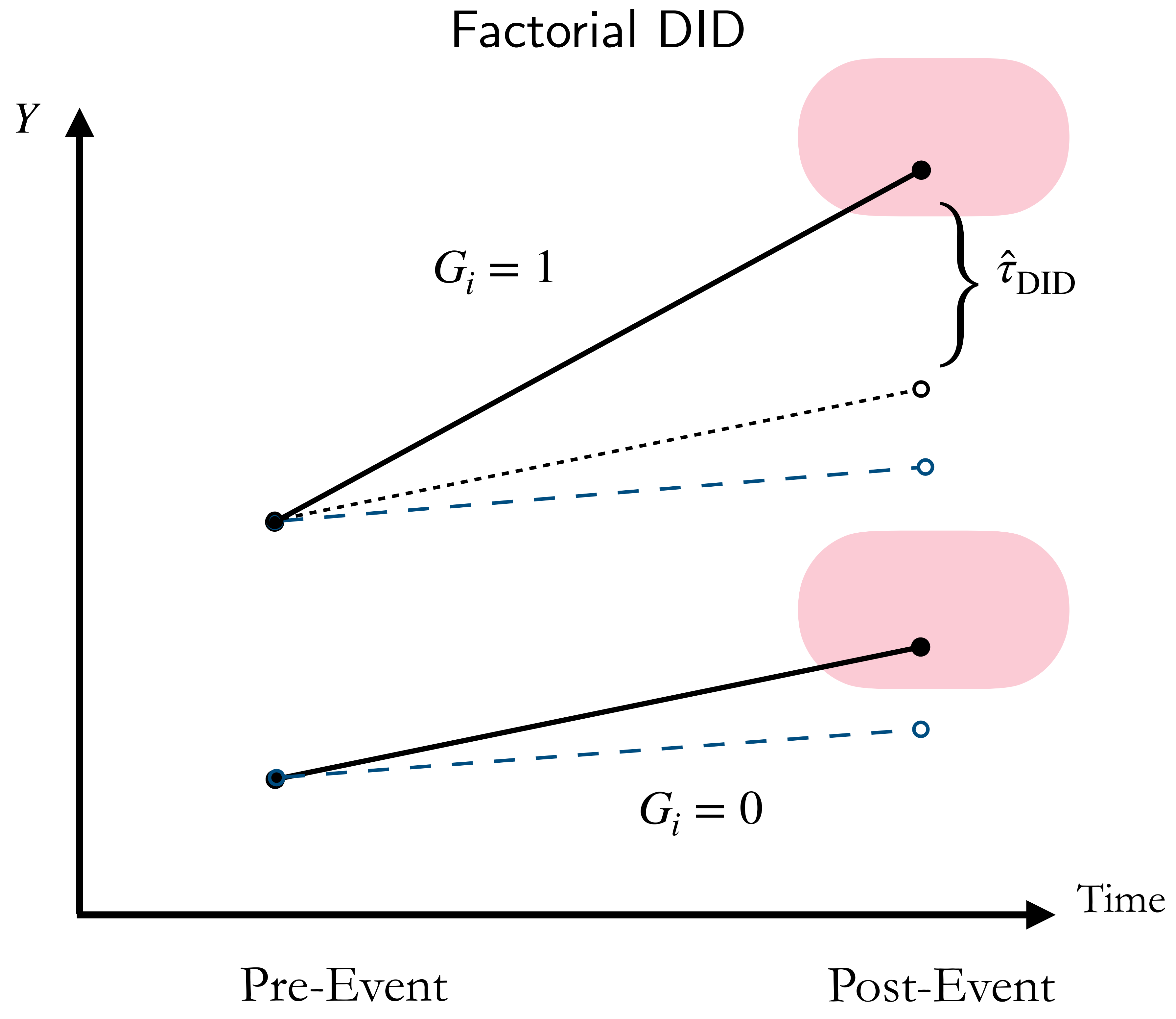


This Paper

Setting, Estimand, Estimator,
ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID

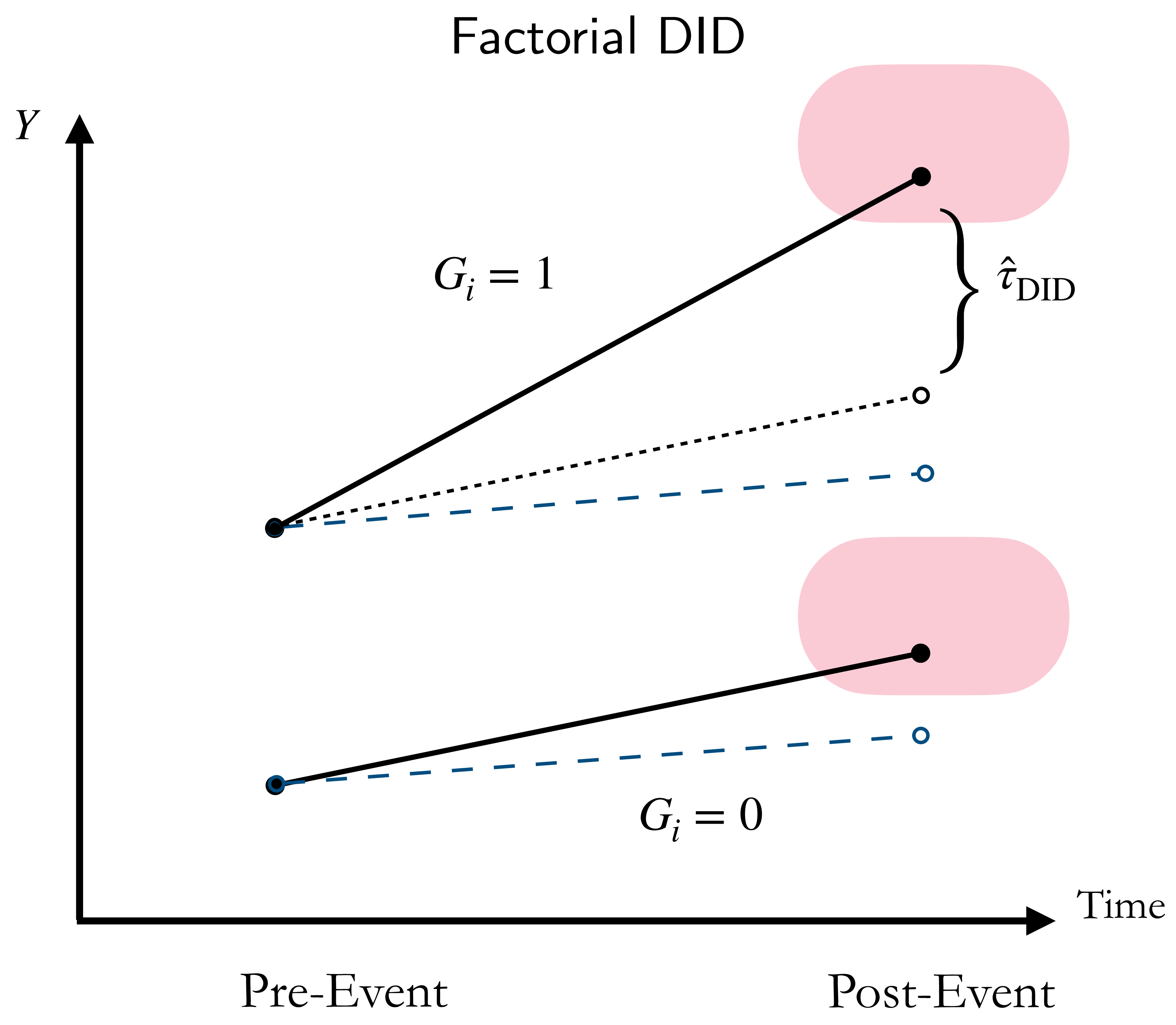
- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event
  - Not causal
  - Need additional analytical tools (factorial designs) to clarify
  - With covariates, common TWFE models need modificaiton



This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

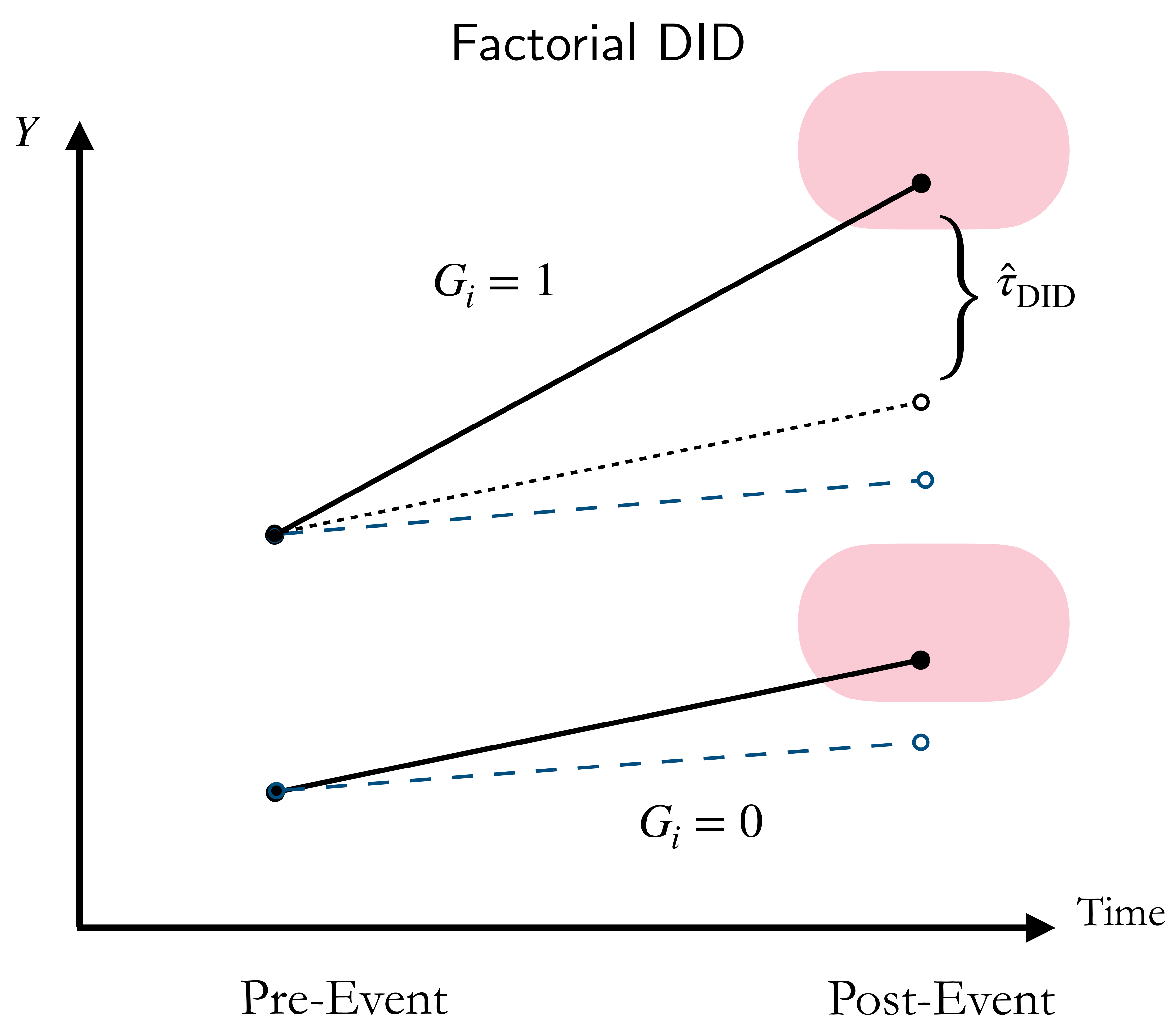
- Factorial DID is a different research design from a canonical DID
- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event
  - Not causal
  - Need additional analytical tools (factorial designs) to clarify
  - With covariates, common TWFE models need modificaiton
- Identifying G's causal effect requires stronger assumptions



This Paper

Setting, Estimand, Estimator, ID Assumptions & ID Results

- Factorial DID is a different research design from a canonical DID
- Under no anticipation & parallel trends, the DID estimator identifies treatment effect heterogeneity of the event
  - Not causal
  - Need additional analytical tools (factorial designs) to clarify
  - With covariates, common TWFE models need modificaiton
- Identifying G's causal effect requires stronger assumptions
- Factorial DID includes canonical DID as a special case with an additional assumption



Related Literature

- DID and TWFE
  - "Regression DD": Card (1992); Angrist & Pischke (2009); Shahn & Hatfield (2024)
  - For reviews of recent development: Roth et al. (2023); Chiu et al. (2023); Arkhangelsky and Imbens (2023)

- Factorial designs
  - VanderWeele (2009); Dasgupta et al (2015); Bansak (2020); Zhao and Ding (2021)

- Bartik instruments & shift shares (e.g., local industry share $\times$ common temporal shock)
  - e.g. Paul Goldsmith-Pinkham et al. (2020); Borusyak, Hull & Jaravel (2022)

- Lord's paradox
  - Lord (1967); Holland and Rubin (1986)

Roadmap

- Motivation
- Setup & Estimands
- Identification
- Extensions
- Example: Clans and Calamity

Roadmap

- Motivation
- Setup & Estimands
- Identification
- Extensions
- Example: Clans and Calamity

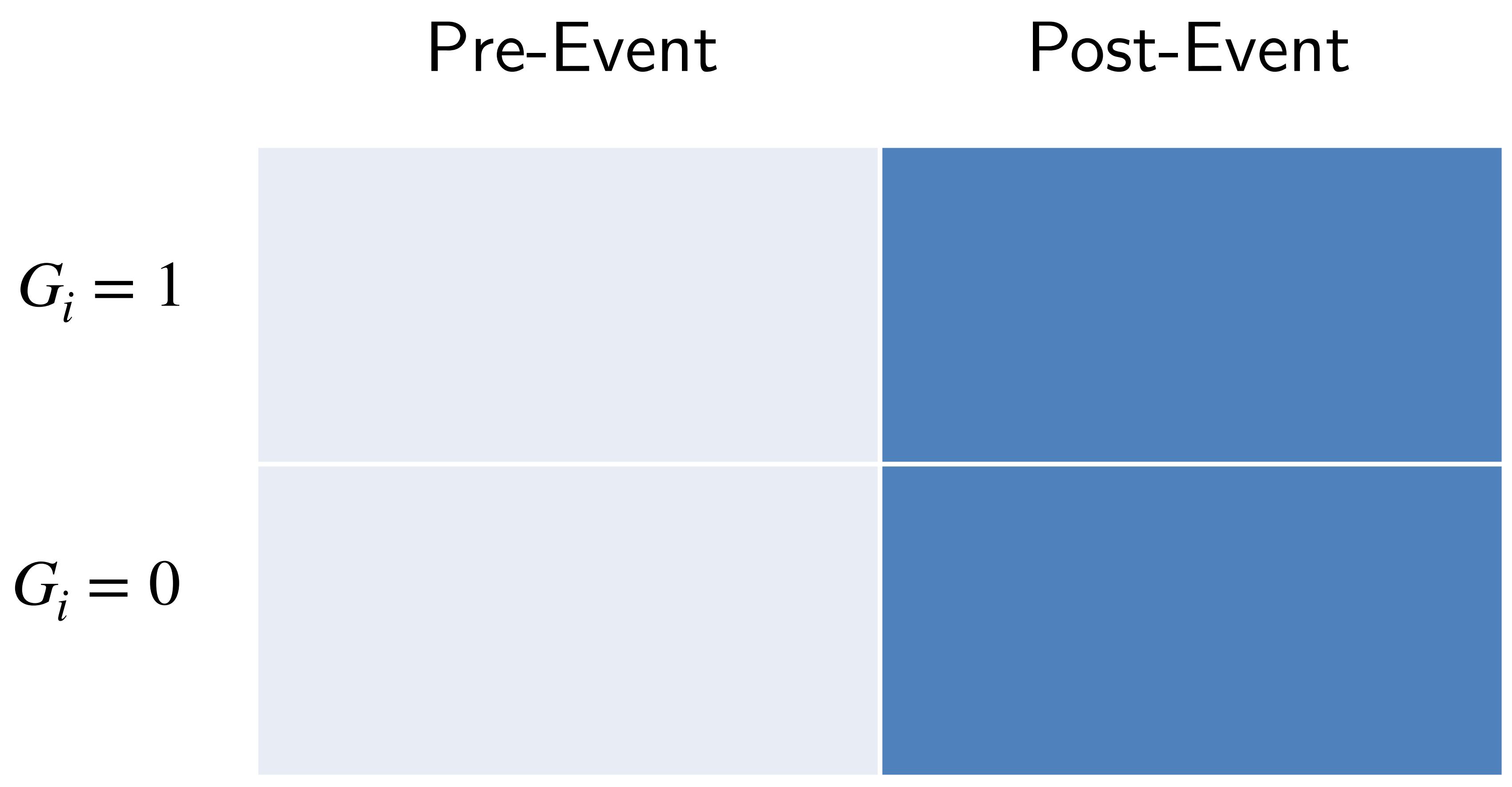
Setup

Setup

Two-group, two-periods; no covariates

Setup

Two-group, two-periods; no covariates

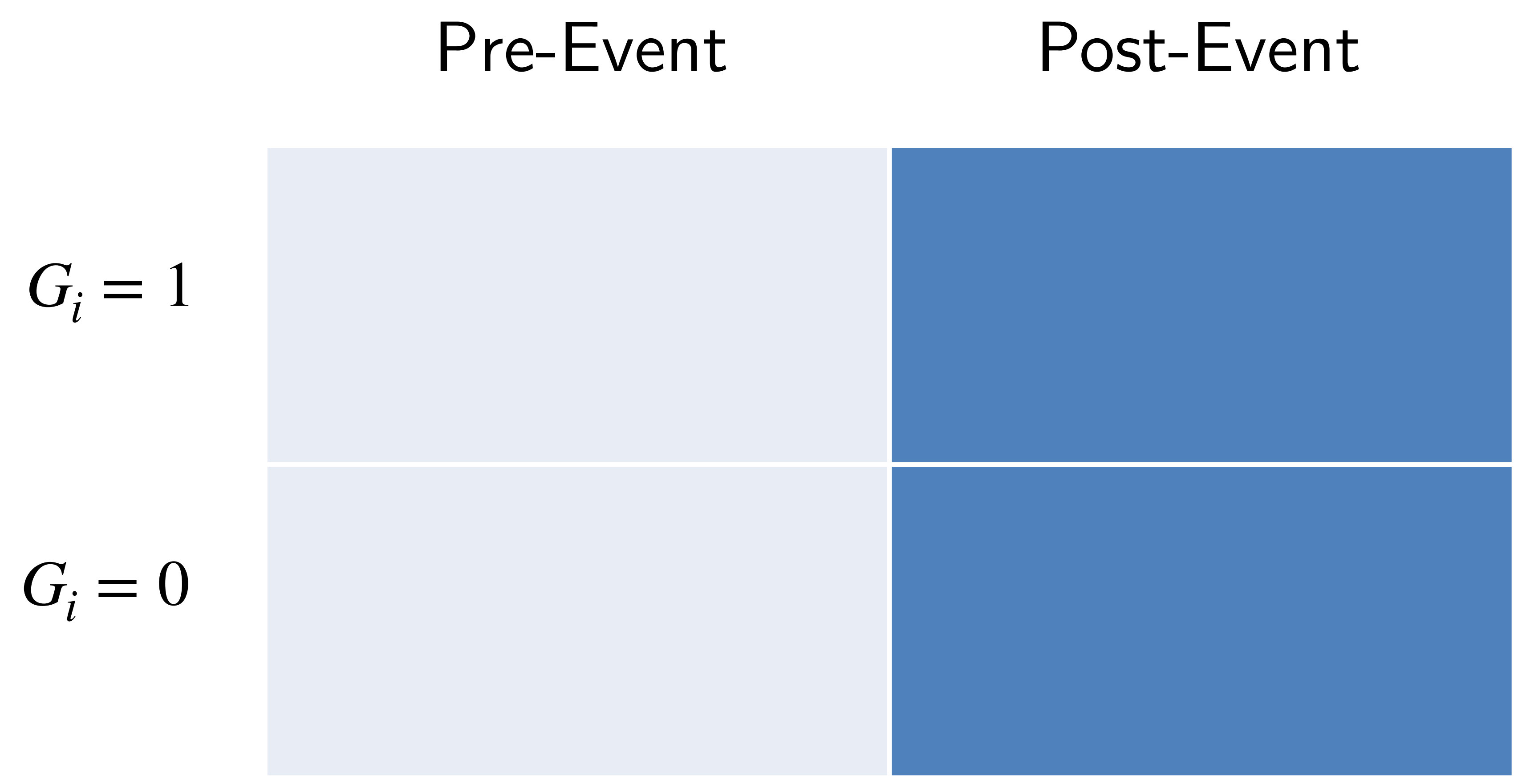


Factorial DID

Setup

Two-group, two-periods; no covariates

- Study population: $i = 1,2,\cdots,n$
- Timing of the event is fixed
- Time periods: $t =$ pre, post
- Baseline factor: $G_i \in \{0,1\}$
- Data: $\{G_i, Y_{i,\text{pre}}, Y_{i,\text{post}} : i = 1,2,\cdots,n\}$

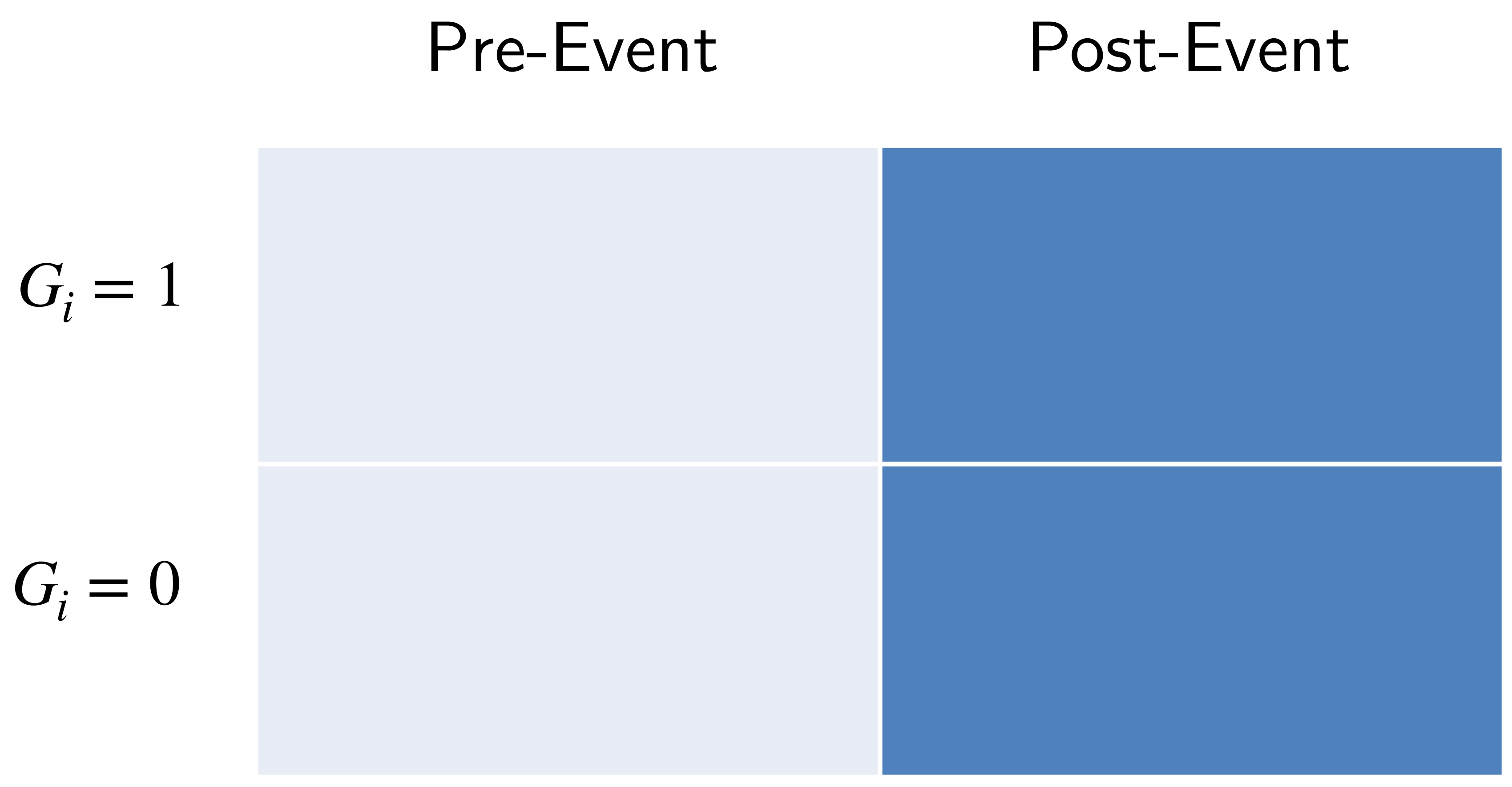


Factorial DID

Setup

Two-group, two-periods; no covariates

- Study population: $i = 1,2,\cdots,n$
- Timing of the event is fixed
- Time periods: $t =$ pre, post
- Baseline factor: $G_i \in \{0,1\}$
- Data: $\{G_i, Y_{i,\text{pre}}, Y_{i,\text{post}} : i = 1,2,\cdots,n\}$
- Exposure to the event in the post period: $Z_i = 1, \;\; i = 1,2,\cdots,n$



Factorial DID

Setup

Two-group, two-periods; no covariates

- Study population: $i = 1,2,\cdots,n$
- Timing of the event is fixed
- Time periods: $t =$ pre, post
- Baseline factor: $G_i \in \{0,1\}$
- Data: $\{G_i, Y_{i,\text{pre}}, Y_{i,\text{post}} : i = 1,2,\cdots,n\}$
- Exposure to the event in the post period: $Z_i = 1, \;\; i = 1,2,\cdots,n$

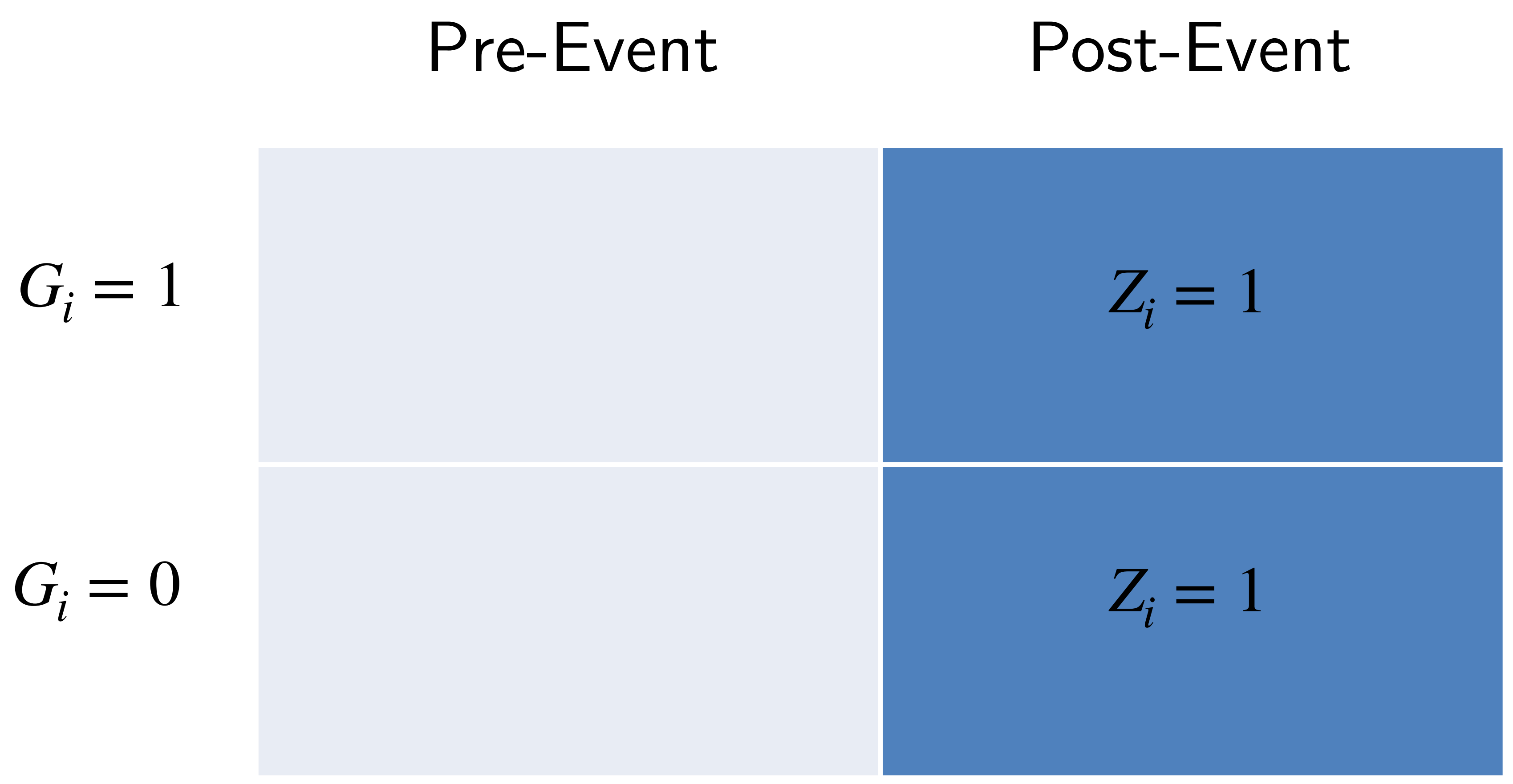
	Pre-Event	Post-Event
$G_i = 1$		$Z_i = 1$
$G_i = 0$		$Z_i = 1$

Factorial DID

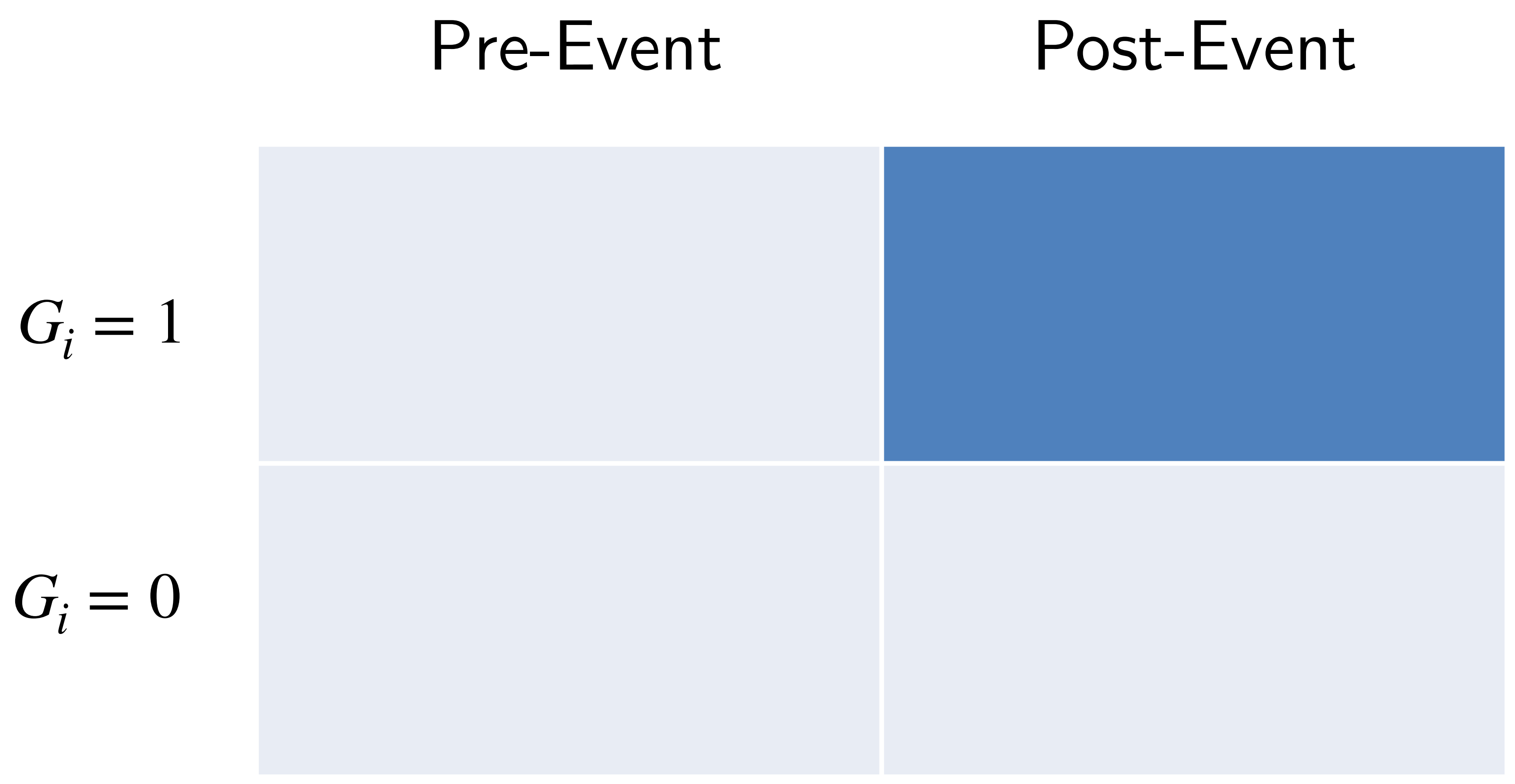
Setup

Two-group, two-periods; no covariates

- Study population: $i = 1,2,\cdots,n$
- Timing of the event is fixed
- Time periods: $t =$ pre, post
- Baseline factor: $G_i \in \{0,1\}$
- Data: $\{G_i, Y_{i,\text{pre}}, Y_{i,\text{post}} : i = 1,2,\cdots,n\}$
- Exposure to the event in the post period: $Z_i = 1, \;\; i = 1,2,\cdots,n$



Factorial DID



Canonical DID

Setup

Two-group, two-periods; no covariates

- Study population: $i = 1,2,\cdots,n$
- Timing of the event is fixed
- Time periods: $t =$ pre, post
- Baseline factor: $G_i \in \{0,1\}$
- Data: $\{G_i, Y_{i,\text{pre}}, Y_{i,\text{post}} : i = 1,2,\cdots,n\}$
- Exposure to the event in the post period: $Z_i = 1, \ \ i = 1,2,\cdots,n$

	Pre-Event	Post-Event
$G_i = 1$		$Z_i = 1$
$G_i = 0$		$Z_i = 1$

Factorial DID

	Pre-Event	Post-Event
$G_i = 1$		$Z_i = 1$
$G_i = 0$		$Z_i = 0$

Canonical DID

The Difference-in-Differences Estimator

	Pre-Event	Post-Event
$G_i = 1$	$\frac{1}{n_1} \sum_{i:G_i=1} Y_{i,\text{pre}}$	$\frac{1}{n_1} \sum_{i:G_i=1} Y_{i,\text{post}}$
$G_i = 0$	$\frac{1}{n_0} \sum_{i:G_i=0} Y_{i,\text{pre}}$	$\frac{1}{n_0} \sum_{i:G_i=0} Y_{i,\text{post}}$

	Pre-Event	Post-Event
$G_i = 1$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$
$G_i = 0$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$

The Difference-in-Differences Estimator

	Pre-Event	Post-Event
$G_i = 1$	$\frac{1}{n_1} \sum_{i:G_i=1} Y_{i,\text{pre}}$	$\frac{1}{n_1} \sum_{i:G_i=1} Y_{i,\text{post}}$
$G_i = 0$	$\frac{1}{n_0} \sum_{i:G_i=0} Y_{i,\text{pre}}$	$\frac{1}{n_0} \sum_{i:G_i=0} Y_{i,\text{post}}$

	Pre-Event	Post-Event
$G_i = 1$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$
$G_i = 0$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$

Define

The Difference-in-Differences Estimator

	Pre-Event	Post-Event
$G_i = 1$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{pre}}$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{post}}$
$G_i = 0$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{pre}}$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{post}}$

	Pre-Event	Post-Event
$G_i = 1$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$
$G_i = 0$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$

Define

- $\hat{\tau}_{\text{DID}} = \frac{1}{n_1}\sum_{i:G_i=1}(Y_{i,\text{post}} - Y_{i,\text{pre}}) - \frac{1}{n_0}\sum_{i:G_i=0}(Y_{i,\text{post}} - Y_{i,\text{pre}})$

The Difference-in-Differences Estimator

	Pre-Event	Post-Event
$G_i = 1$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{pre}}$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{post}}$
$G_i = 0$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{pre}}$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{post}}$

	Pre-Event	Post-Event
$G_i = 1$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$
$G_i = 0$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$

Define

- $\hat{\tau}_{\text{DID}} = \frac{1}{n_1}\sum_{i:G_i=1}(Y_{i,\text{post}} - Y_{i,\text{pre}}) - \frac{1}{n_0}\sum_{i:G_i=0}(Y_{i,\text{post}} - Y_{i,\text{pre}})$
- Focus on $\tau_{\text{DID}} = \text{plim } \hat{\tau}_{\text{DID}}$

The Difference-in-Differences Estimator

	Pre-Event	Post-Event
$G_i = 1$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{pre}}$	$\frac{1}{n_1}\sum_{i:G_i=1} Y_{i,\text{post}}$
$G_i = 0$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{pre}}$	$\frac{1}{n_0}\sum_{i:G_i=0} Y_{i,\text{post}}$

	Pre-Event	Post-Event
$G_i = 1$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$
$G_i = 0$	$Y_{i,\text{pre}}$	$Y_{i,\text{post}}$

Define

- $\hat{\tau}_{\text{DID}} = \frac{1}{n_1}\sum_{i:G_i=1}(Y_{i,\text{post}} - Y_{i,\text{pre}}) - \frac{1}{n_0}\sum_{i:G_i=0}(Y_{i,\text{post}} - Y_{i,\text{pre}})$

- Focus on $\tau_{\text{DID}} = \text{plim}\ \hat{\tau}_{\text{DID}}$ $= \mathbb{E}[\Delta Y_i \mid G_i = 1] - \mathbb{E}[\Delta Y_i \mid G_i = 0]$

Potential Outcomes

Potential Outcomes

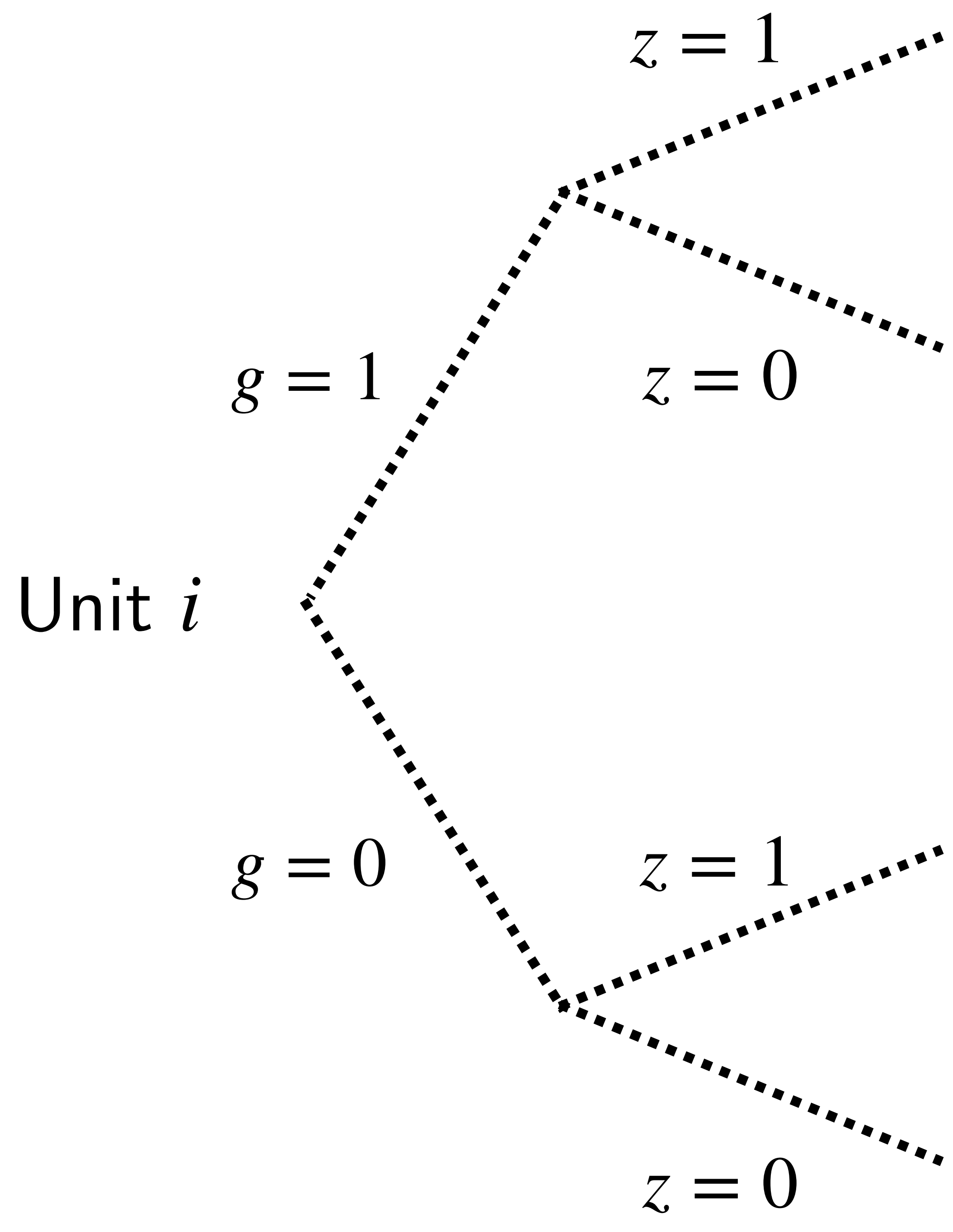
- Frame factorial DID as a factorial design with two factors: g and z

Potential Outcomes

- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g, z)$

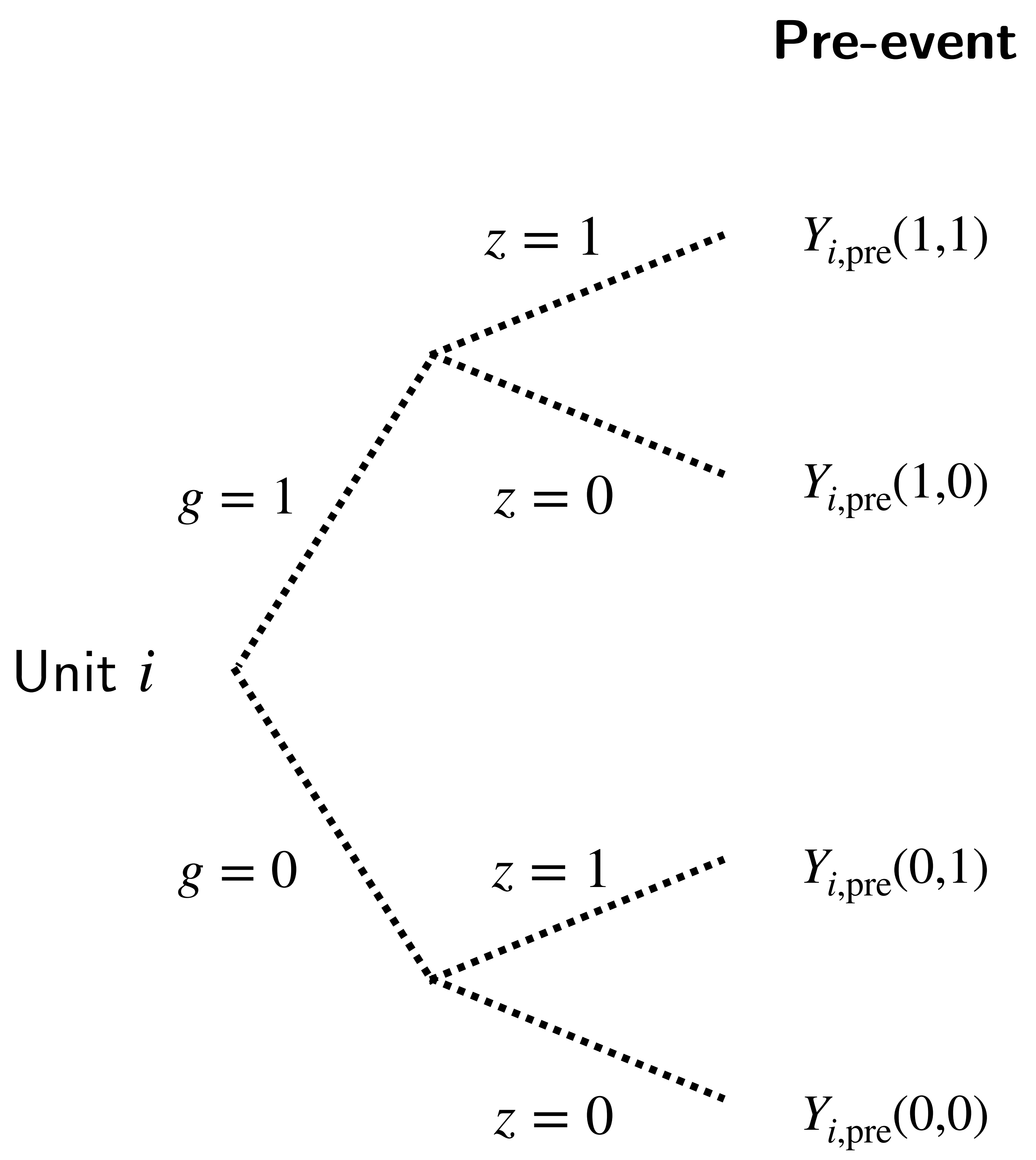
Potential Outcomes

- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g, z)$



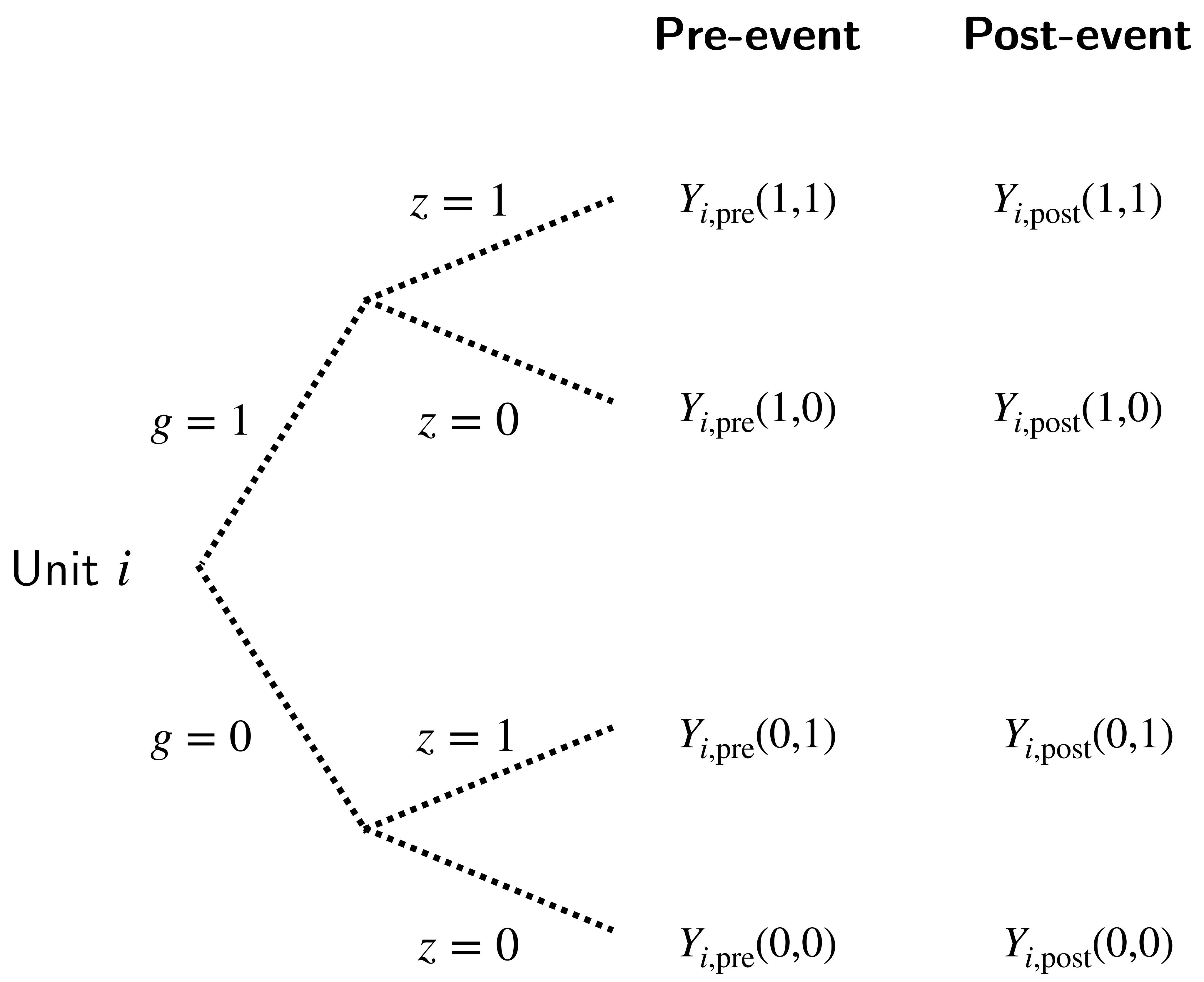
Potential Outcomes

- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g, z)$



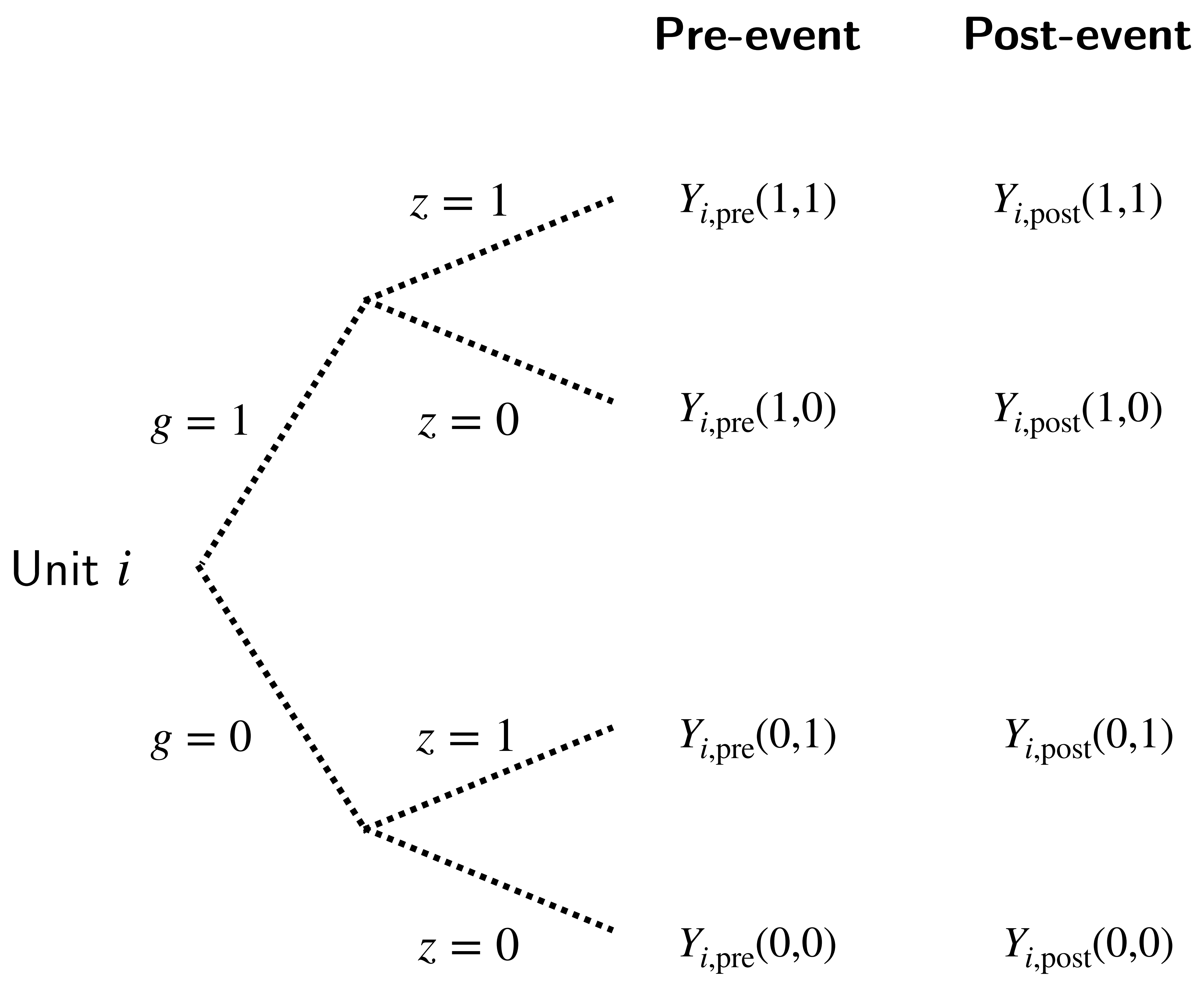
Potential Outcomes

- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g, z)$



Potential Outcomes

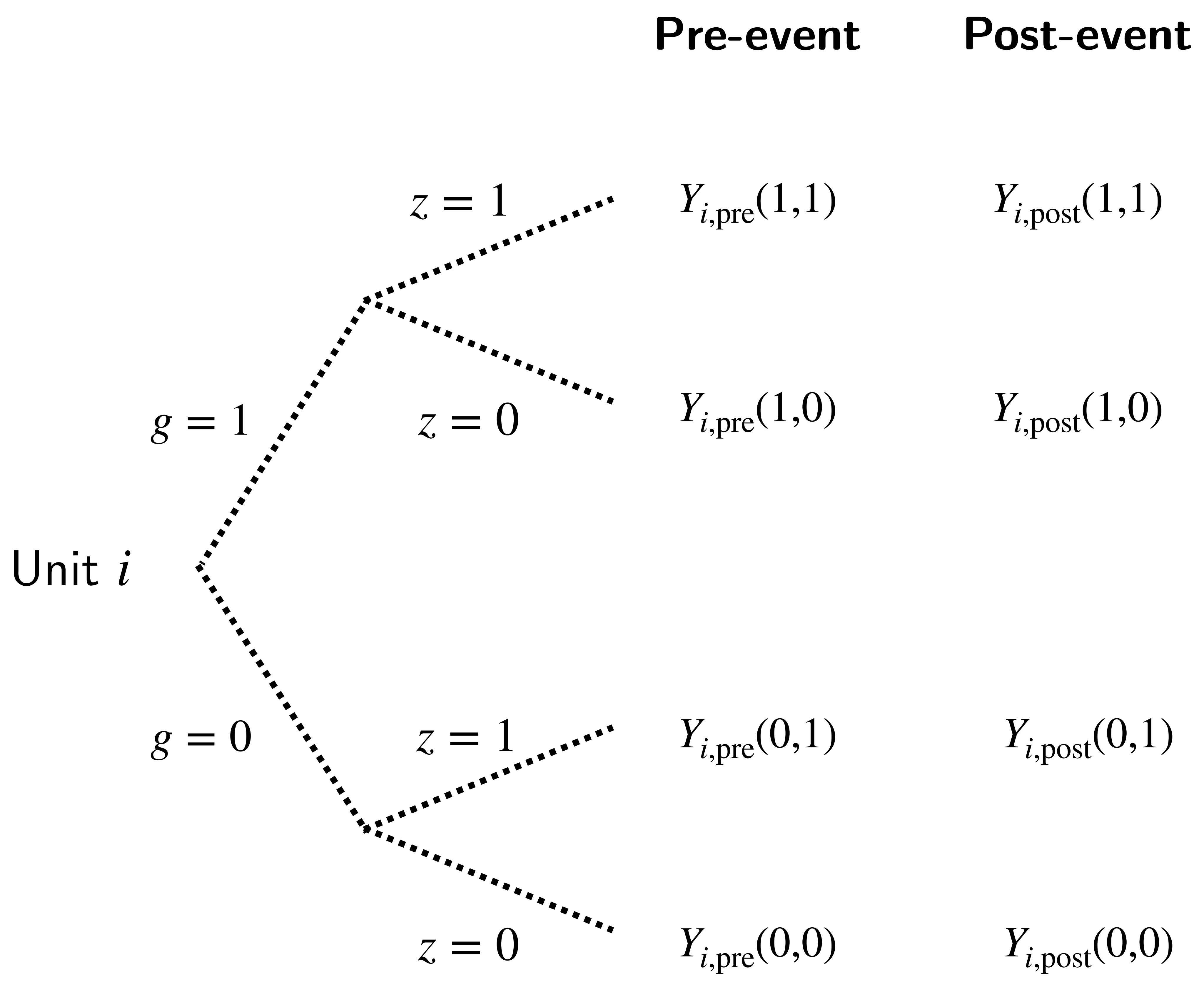
- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g, z)$
- Observed outcomes:



Potential Outcomes

- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g,z)$
- Observed outcomes:

$$Y_{i,t} = \sum_{g=0,1,z=0,1} \mathbf{1}\{G_i = g, Z_i = z\} \cdot Y_{i,t}(g,z) = Y_{i,t}(G_i, Z_i)$$

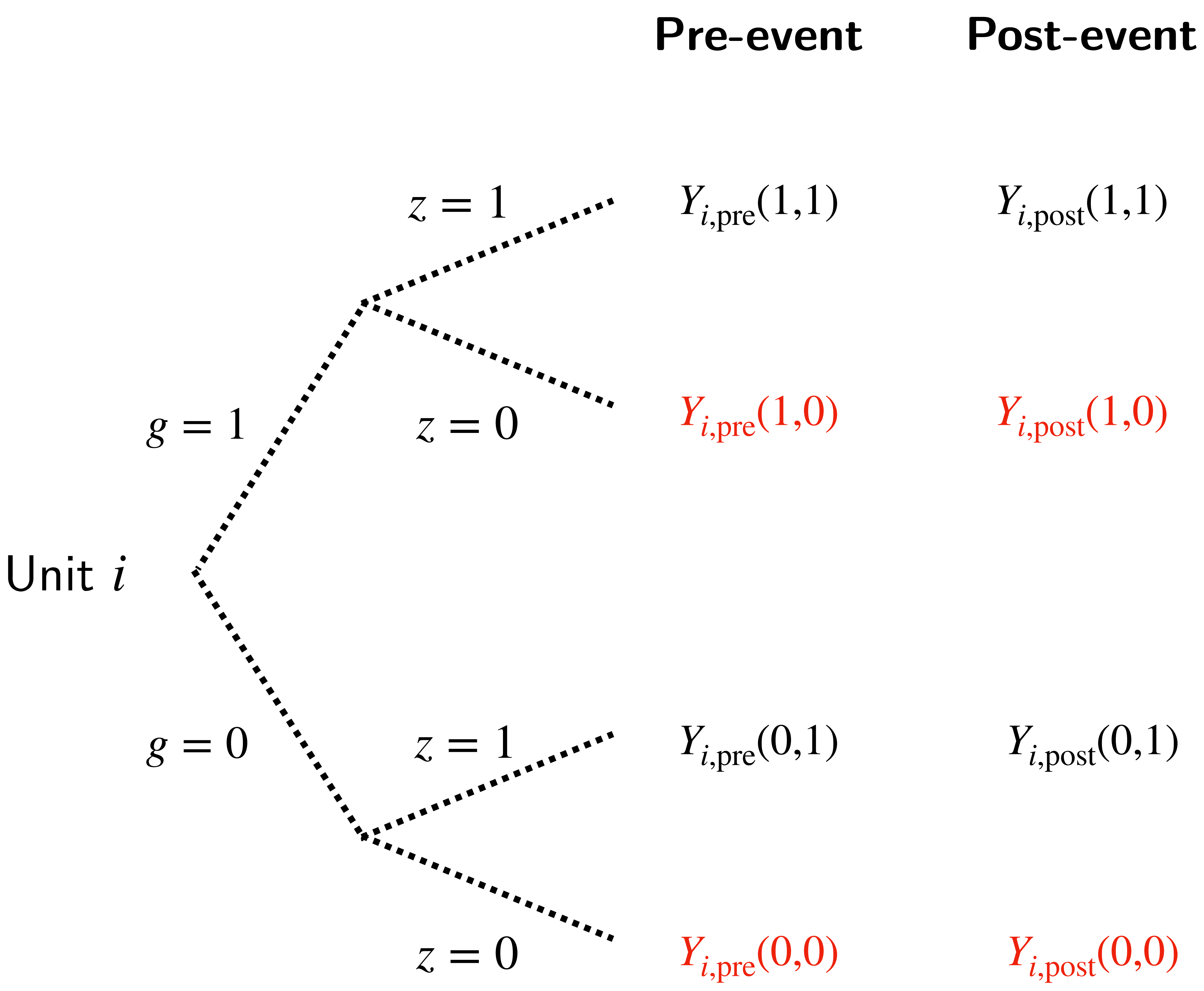


Potential Outcomes

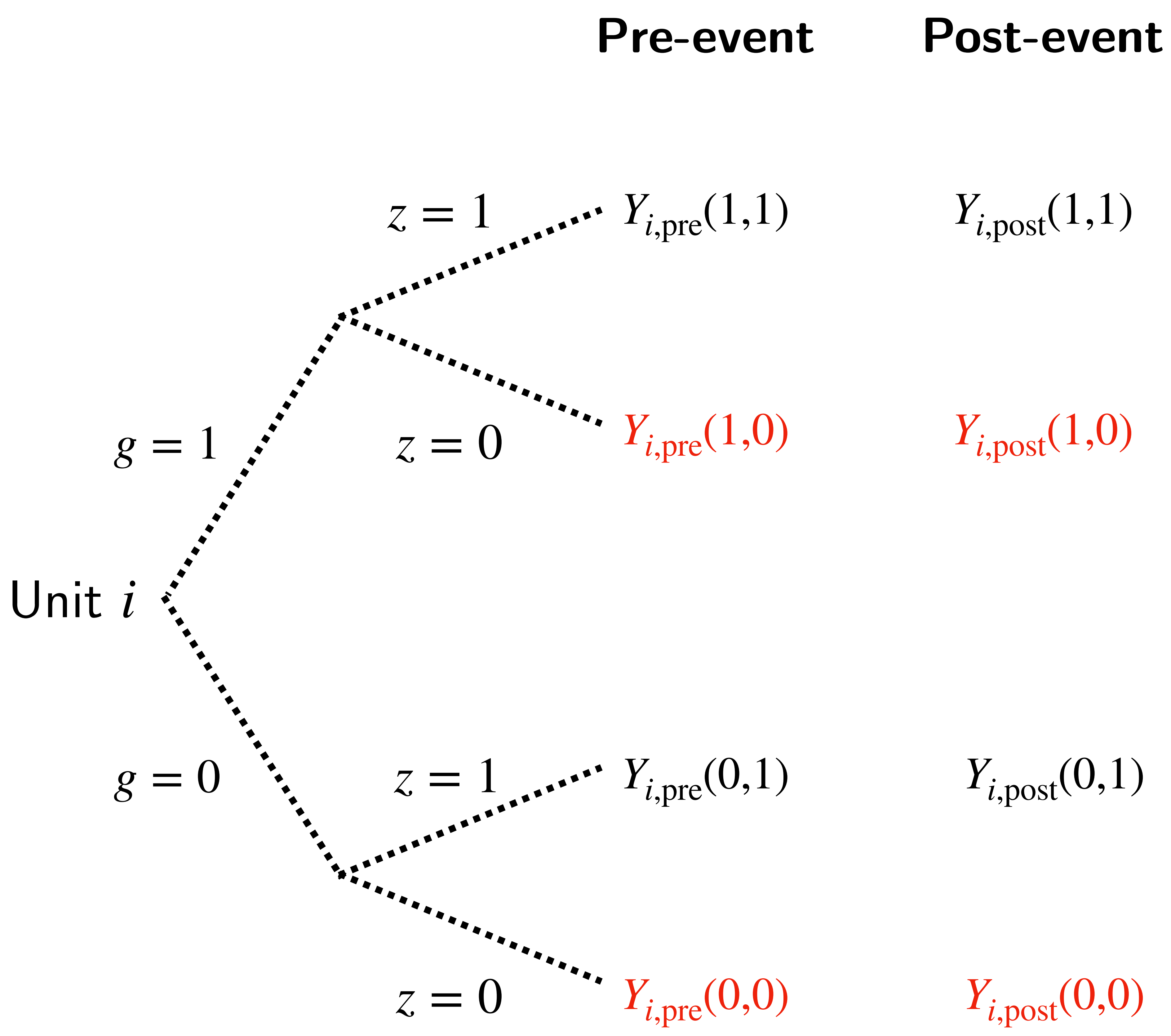
- Frame factorial DID as a factorial design with two factors: g and z
- Potential outcomes indexed by two factors: $Y_{i,t}(g,z)$
- Observed outcomes:

$$Y_{i,t} = \sum_{g=0,1,z=0,1} \mathbf{1}\{G_i = g, Z_i = z\} \cdot Y_{i,t}(g,z) = Y_{i,t}(G_i, Z_i)$$

- Recall: in observed data, $Z_i = 1$ for all units

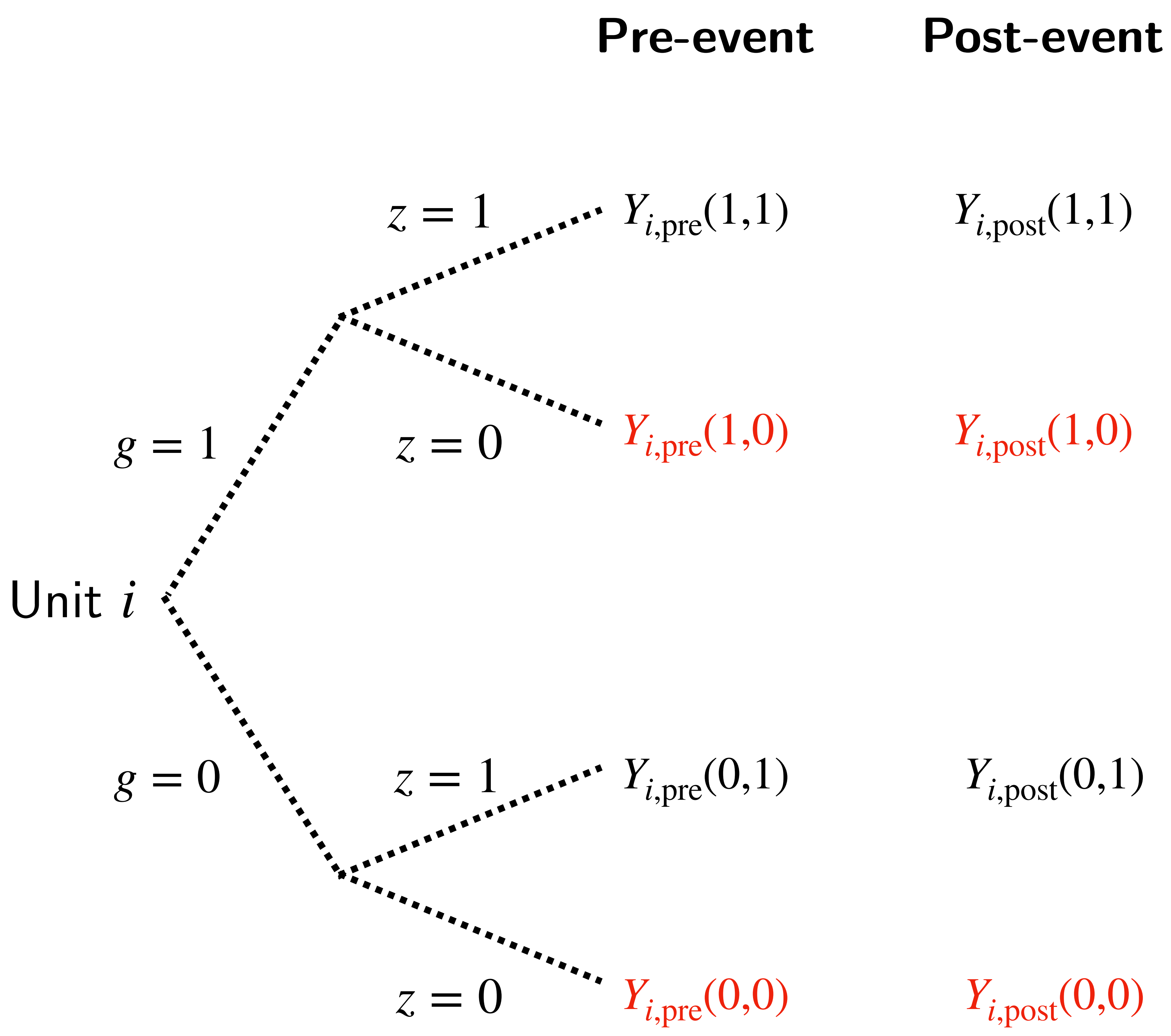


Causal Quantities of Interest



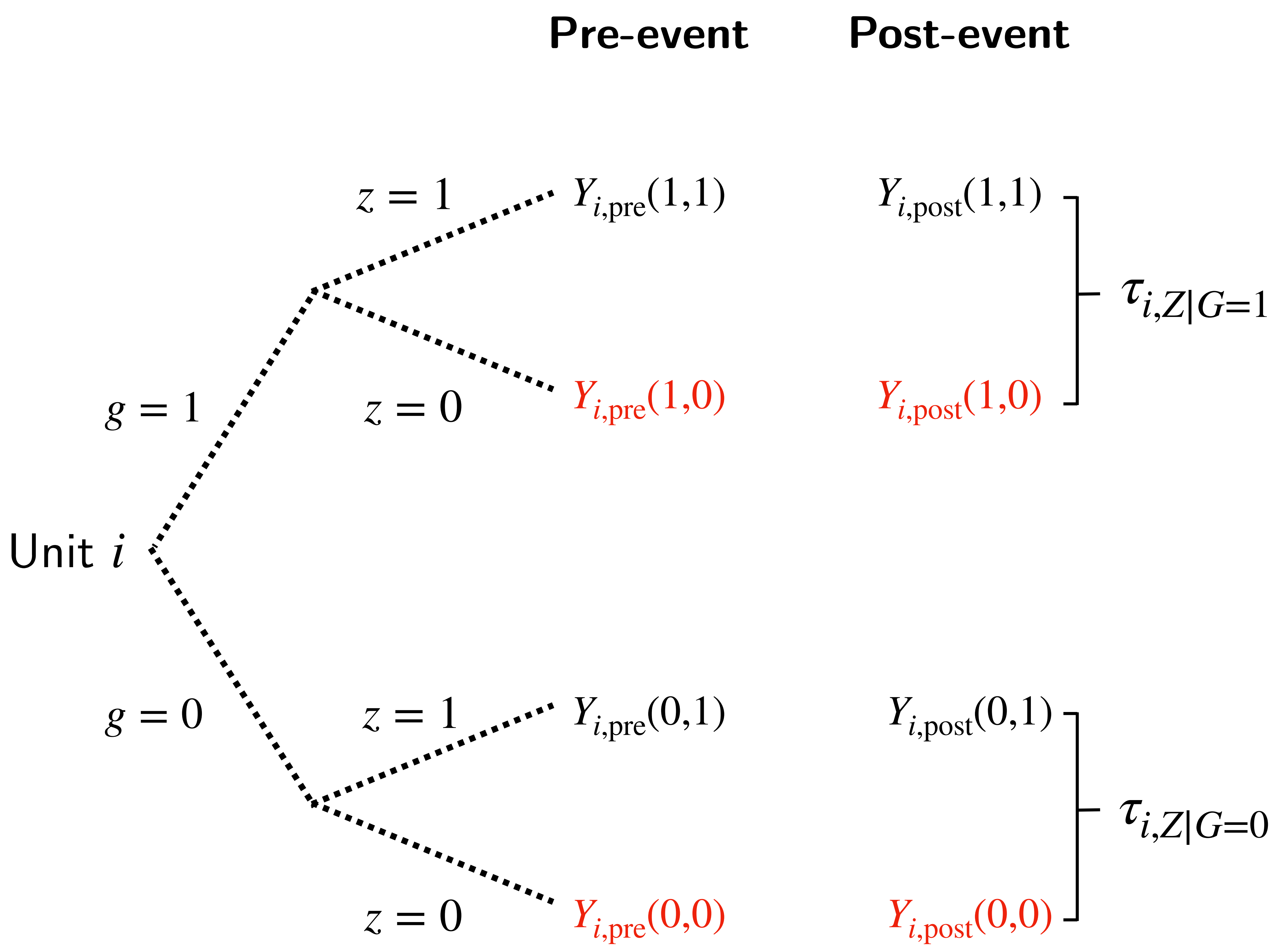
Causal Quantities of Interest

- Individual conditional effect
  - $\tau_{i,Z|G=g} = Y_{i,\text{post}}(g,1) - Y_{i,\text{post}}(g,0)$



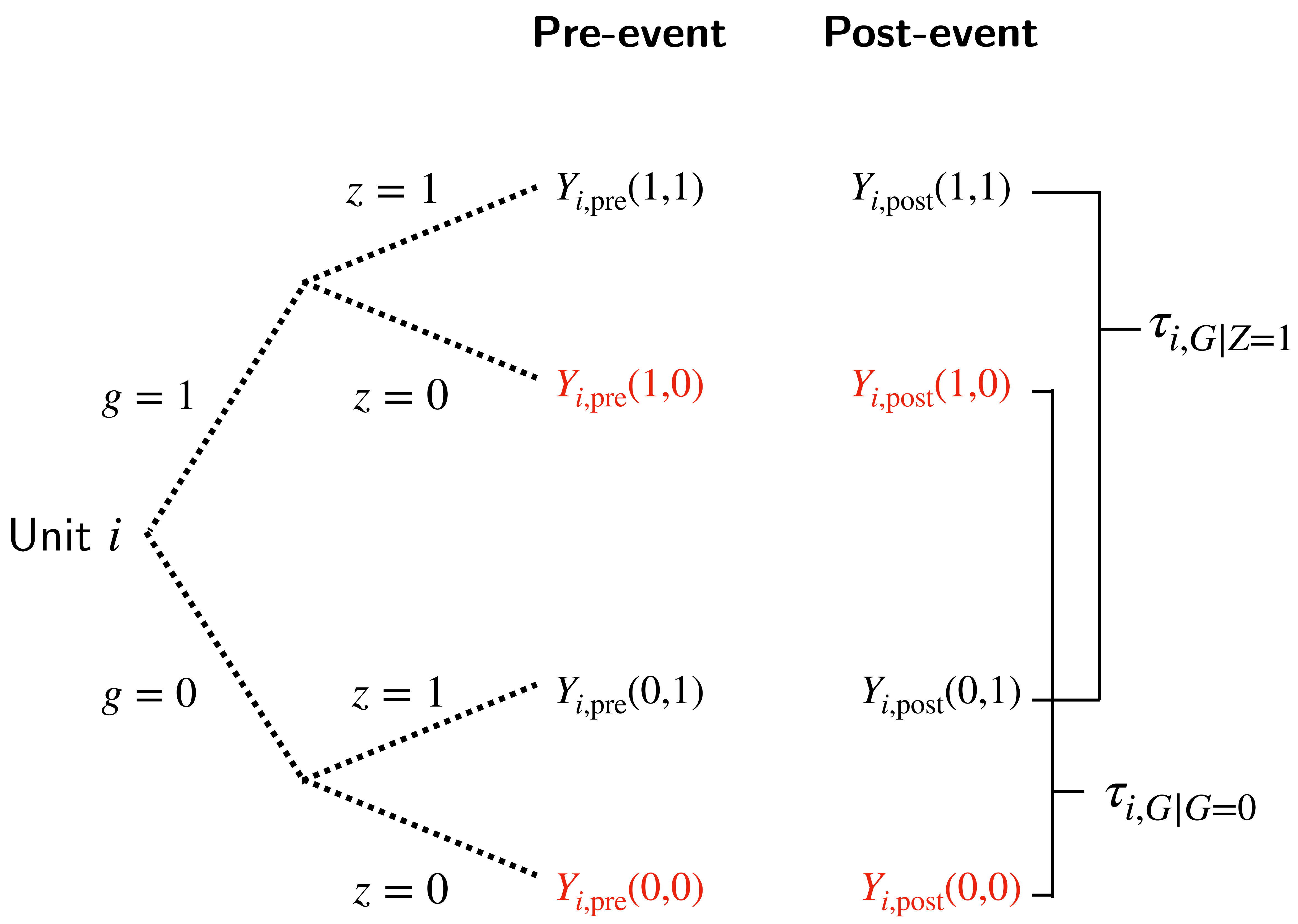
Causal Quantities of Interest

- Individual conditional effect
  - $\tau_{i,Z|G=g} = Y_{i,\text{post}}(g,1) - Y_{i,\text{post}}(g,0)$



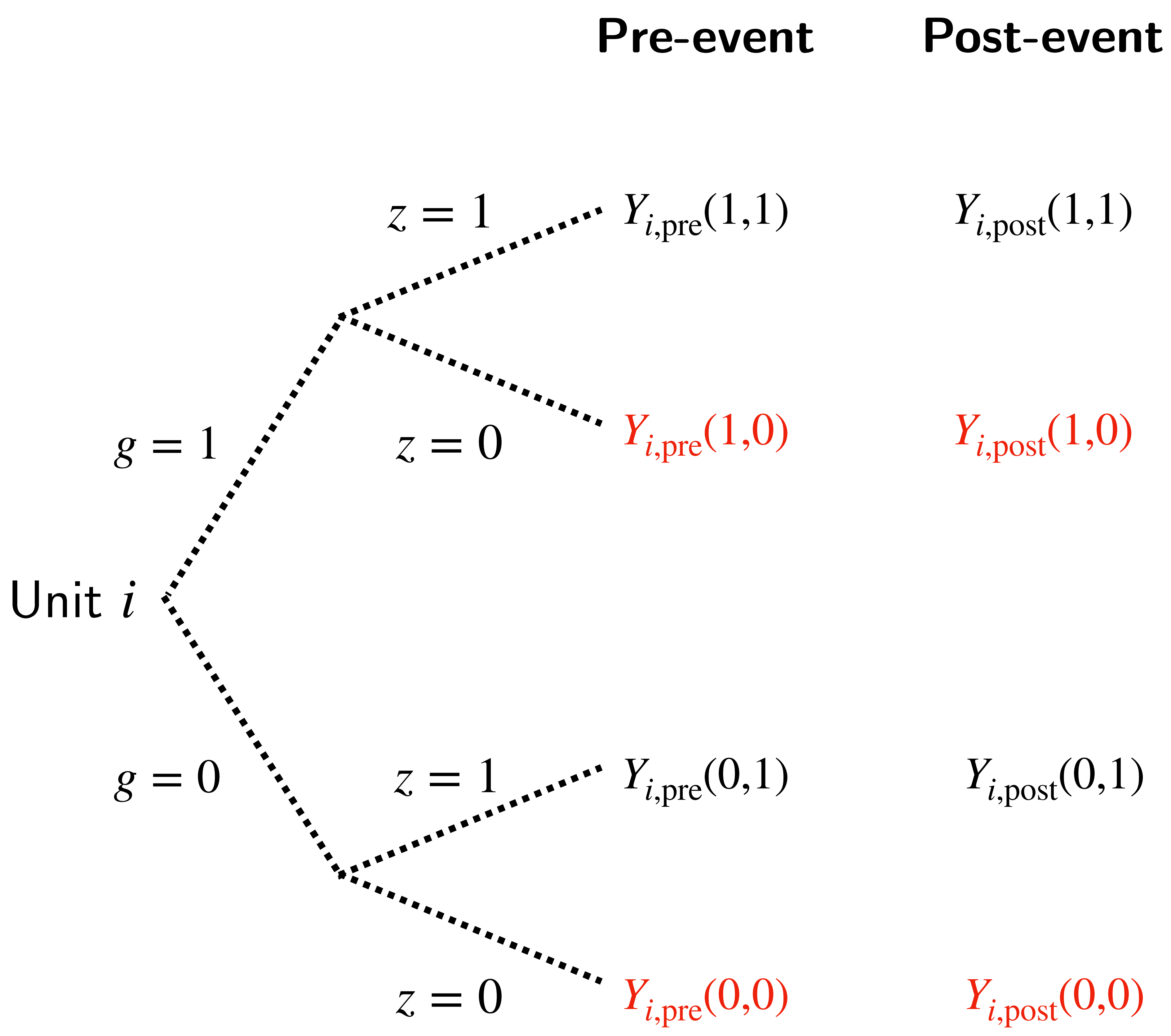
Causal Quantities of Interest

- Individual conditional effect
  - $\tau_{i,Z|G=g} = Y_{i,\text{post}}(g,1) - Y_{i,\text{post}}(g,0)$
  - $\tau_{i,G|Z=z} = Y_{i,\text{post}}(1,z) - Y_{i,\text{post}}(0,z)$



Causal Quantities of Interest

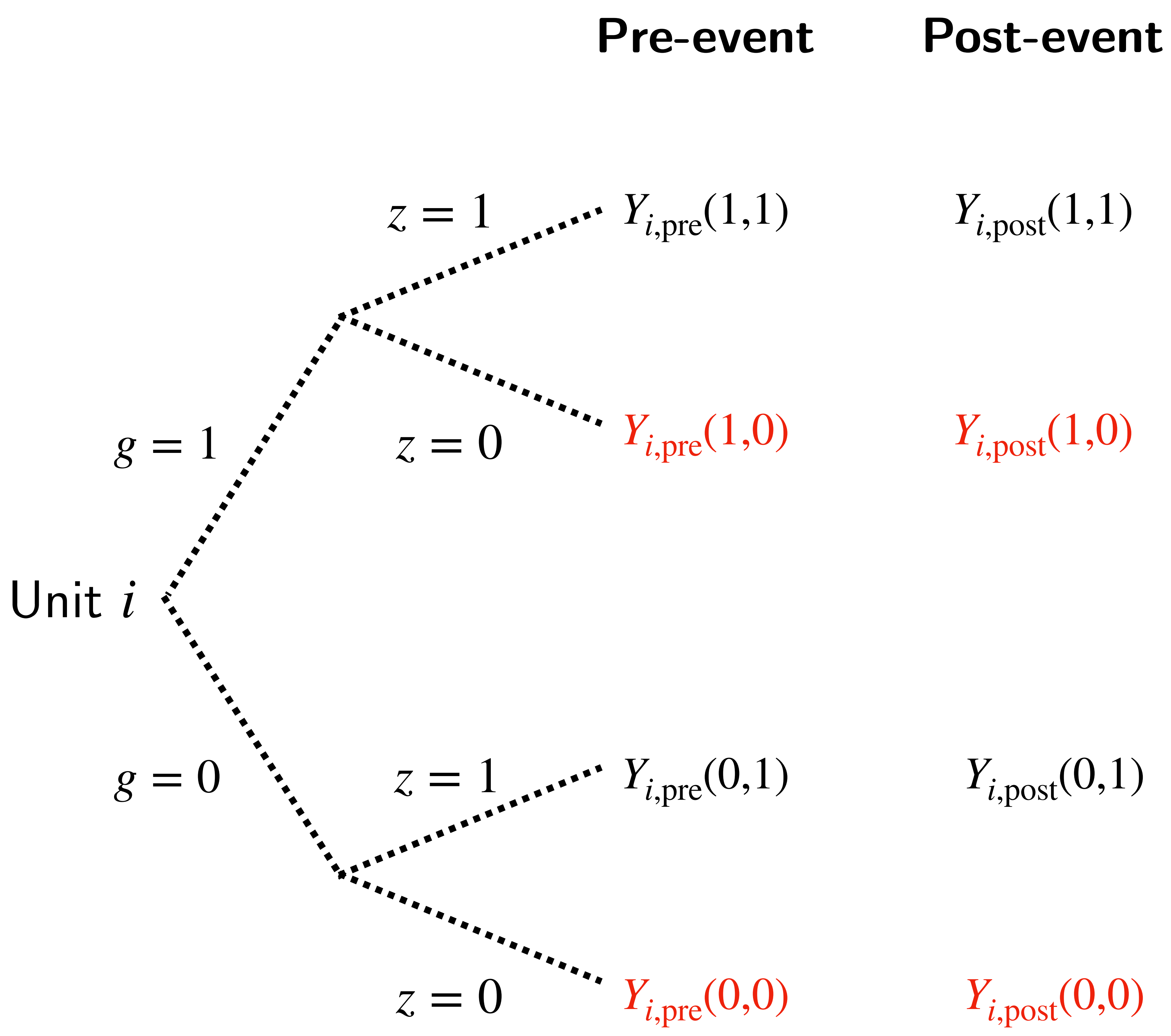
- Individual interaction effect
  - $\tau_{i,\text{inter}} = Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)$



Causal Quantities of Interest

- Individual interaction effect
  - $\tau_{i,\text{inter}} = Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)$

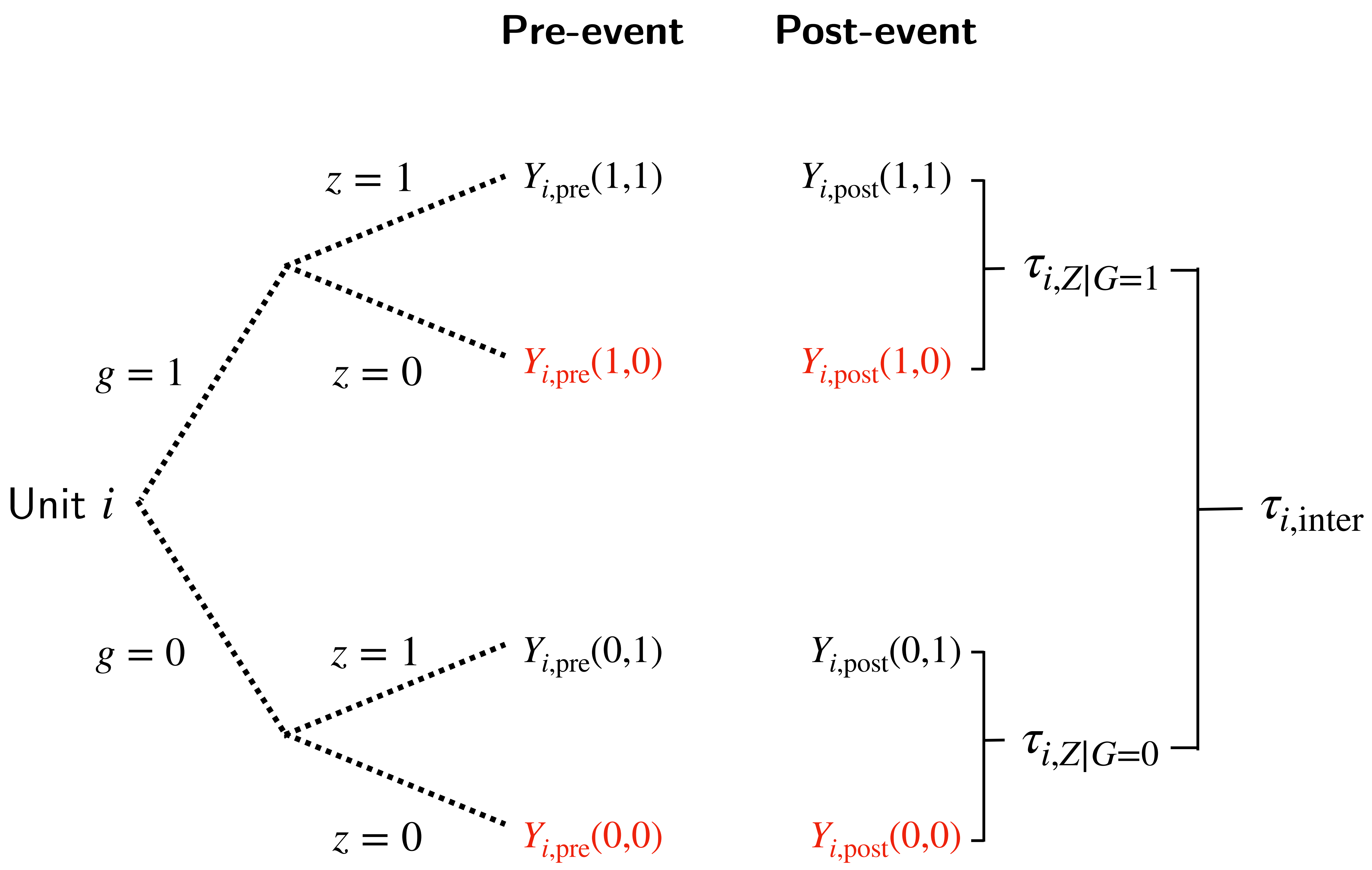
    $= \tau_{i,Z|G=1} - \tau_{i,Z|G=0}$ (Causal moderation of G on Z)



Causal Quantities of Interest

- Individual interaction effect
  - $\tau_{i,\text{inter}} = Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)$

    $= \tau_{i,Z|G=1} - \tau_{i,Z|G=0}$ (Causal moderation of G on Z)

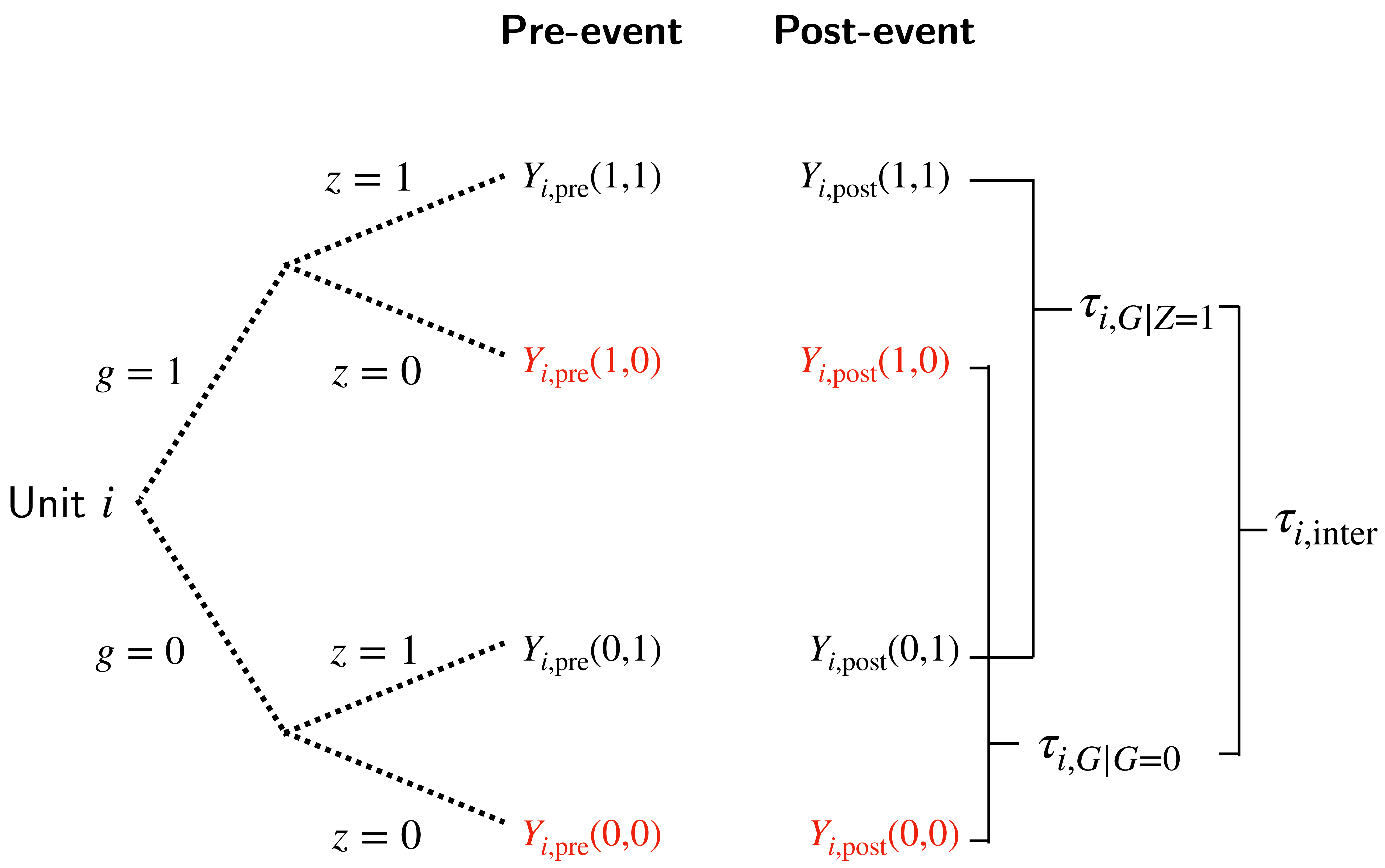


Causal Quantities of Interest

- Individual interaction effect
  - $\tau_{i,\text{inter}} = Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)$

    $= \tau_{i,Z|G=1} - \tau_{i,Z|G=0}$ (Causal moderation of G on Z)

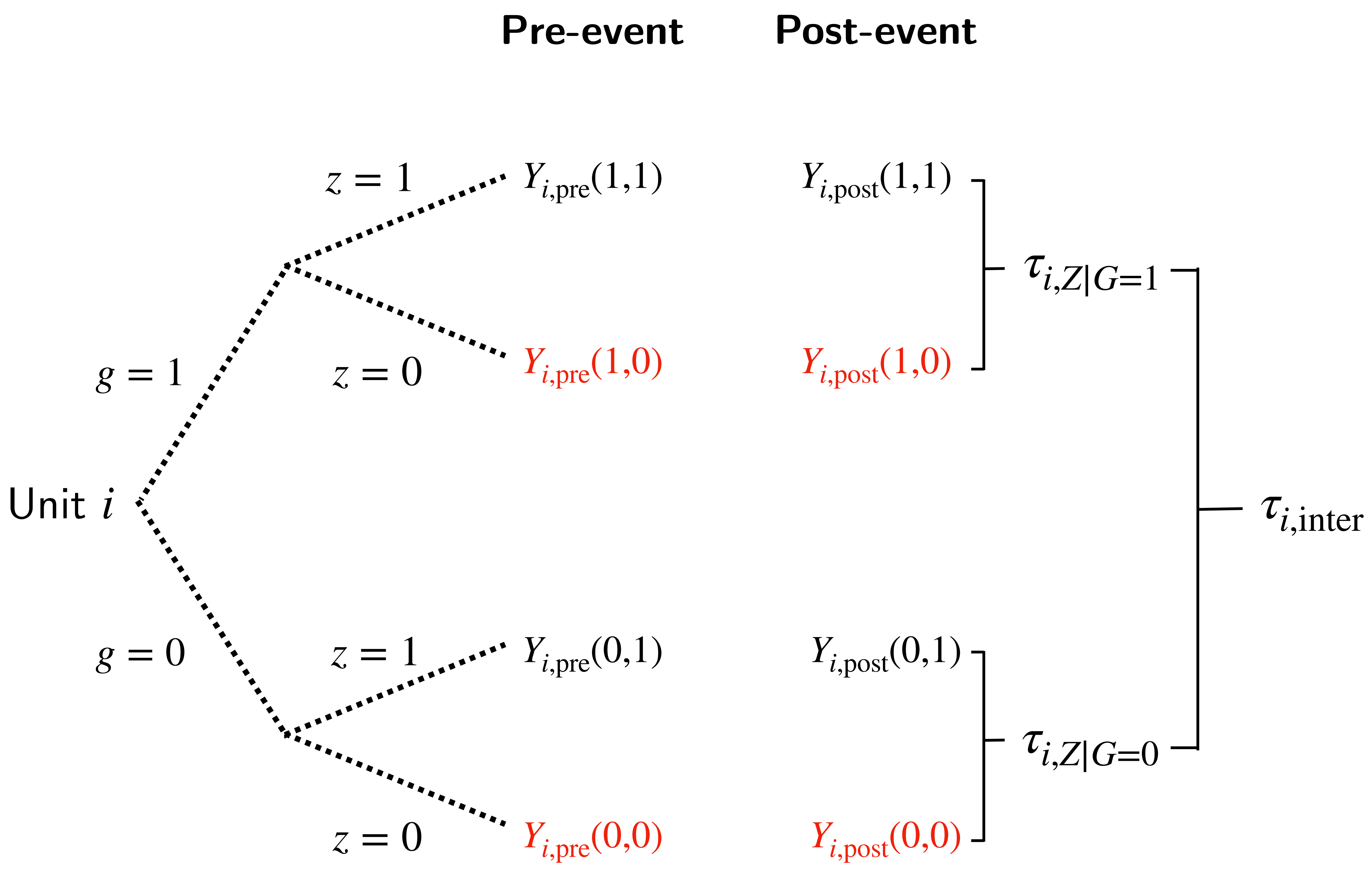
    $= \tau_{i,G|Z=1} - \tau_{i,G|Z=0}$ (Causal moderation of Z on G)



Causal Quantities of Interest

- Individual interaction effect
  - $\tau_{i,\text{inter}} = Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)$

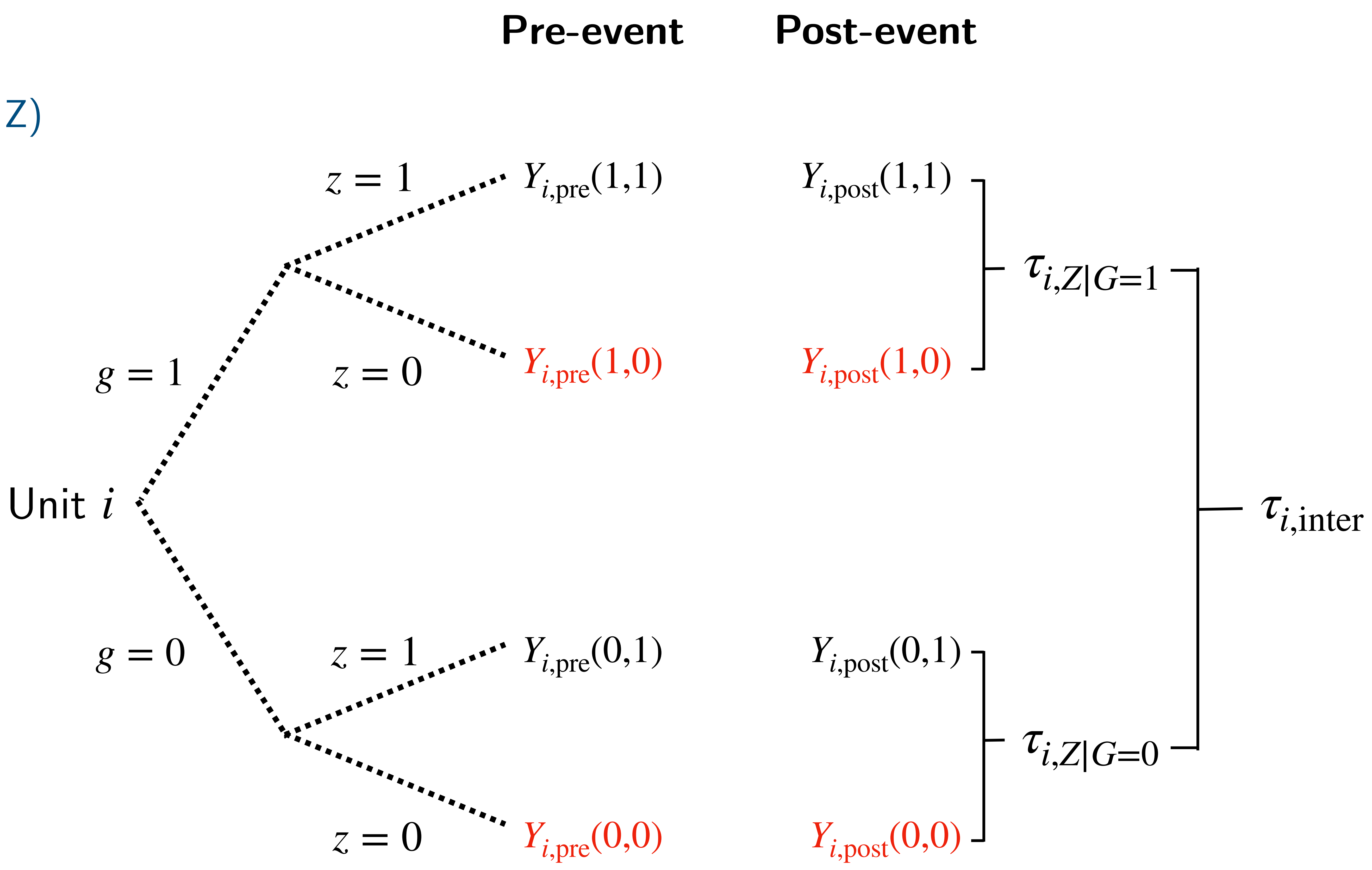
    $= \tau_{i,Z|G=1} - \tau_{i,Z|G=0}$ (Causal moderation of G on Z)



Causal Quantities of Interest

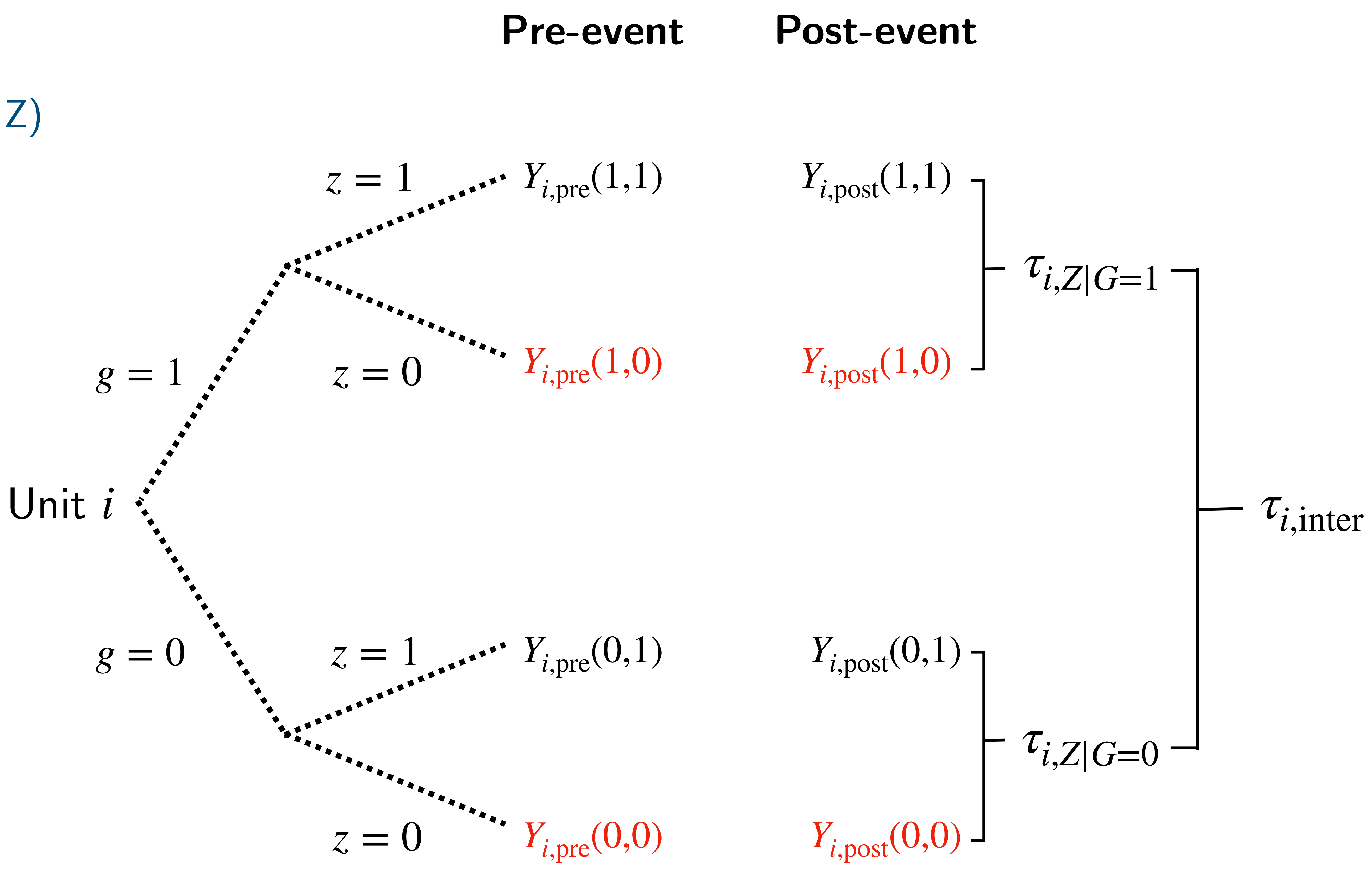
- Average causal interaction (VanderWeele, 2009; Bansak, 2020)
  - $\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)]$

    $= \mathbb{E}[\tau_{i,Z|G=1}] - \mathbb{E}[\tau_{i,Z|G=0}]$ (Causal moderation of G on Z)



Causal Quantities of Interest

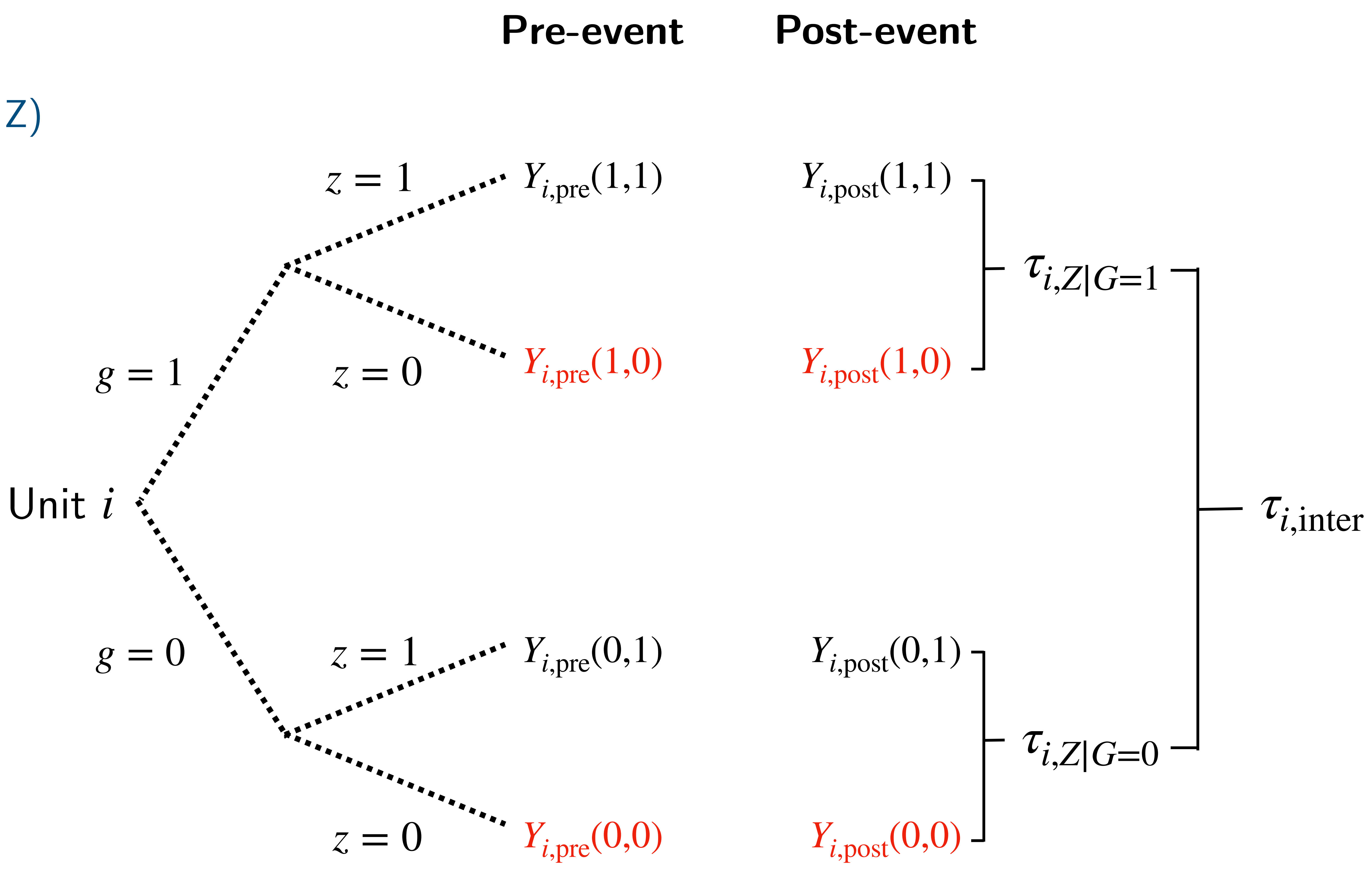
- Average causal interaction (VanderWeele, 2009; Bansak, 2020)
  - $\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)]$
    $= \mathbb{E}[\tau_{i,Z|G=1}] - \mathbb{E}[\tau_{i,Z|G=0}]$ (Causal moderation of G on Z)



Causal Quantities of Interest

- Average causal interaction (VanderWeele, 2009; Bansak, 2020)
  - $\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)]$
    $= \mathbb{E}[\tau_{i,Z|G=1}] - \mathbb{E}[\tau_{i,Z|G=0}]$ (Causal moderation of G on Z)

- Effect modification — Associative
  - $\tau_{\text{em}} = \mathbb{E}[\tau_{i,Z|G=1} \mid G_i = 1] - \mathbb{E}[\tau_{i,Z|G=0} \mid G_i = 0]$

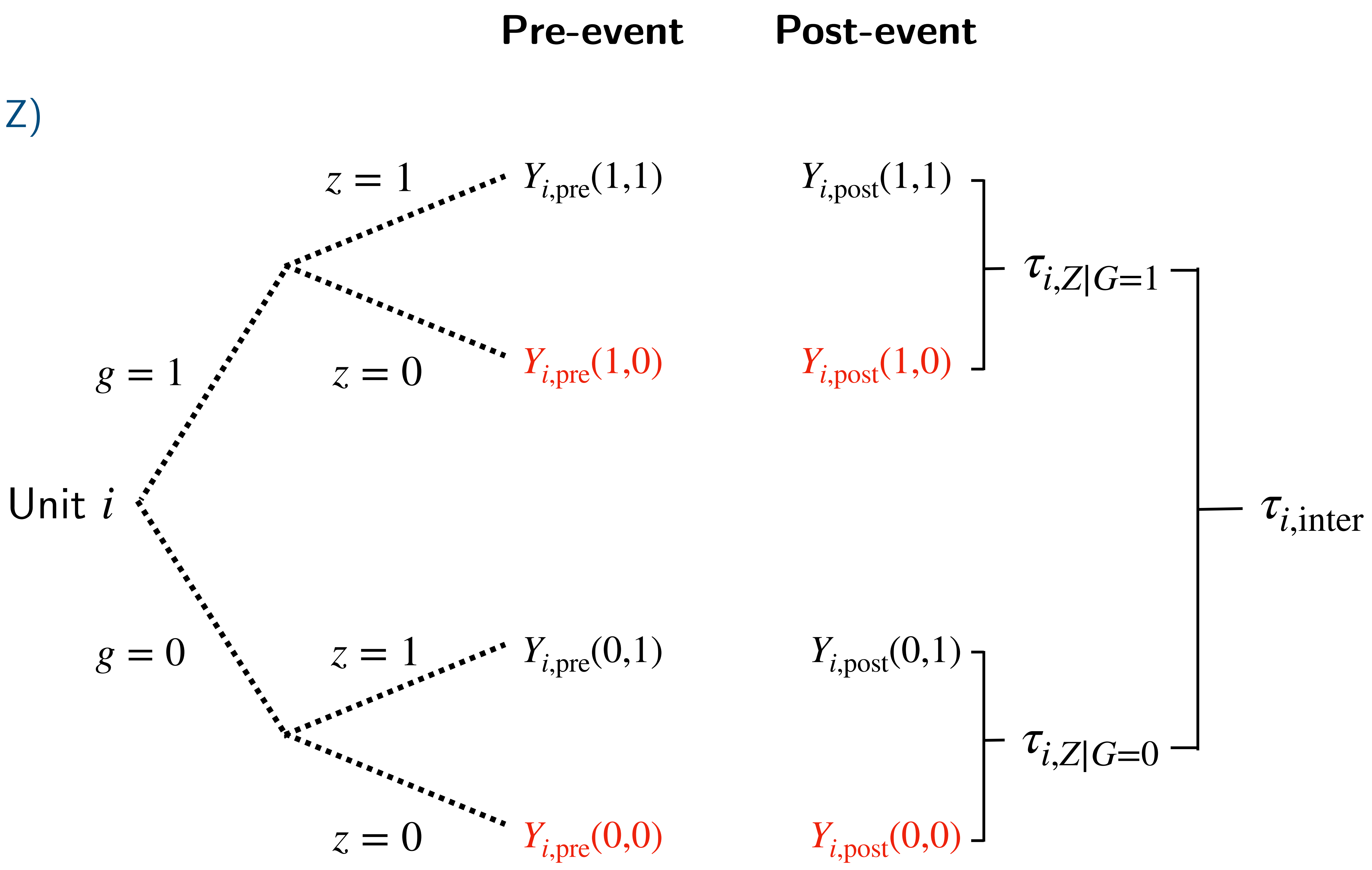


Causal Quantities of Interest

- Average causal interaction (VanderWeele, 2009; Bansak, 2020)
  - $\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) - Y_{i,\text{post}}(0,1) + Y_{i,\text{post}}(0,0)]$

    $= \mathbb{E}[\tau_{i,Z|G=1}] - \mathbb{E}[\tau_{i,Z|G=0}]$ (Causal moderation of G on Z)

- Effect modification — Associative
  - $\tau_{\text{em}} = \mathbb{E}[\tau_{i,Z|G=1} \mid G_i = 1] - \mathbb{E}[\tau_{i,Z|G=0} \mid G_i = 0]$



Roadmap

- Motivation
- Setup & Estimands
- Identification
- Extensions
- Example: Clans and Calamity

Identification: A Roadmap

Identification: A Roadmap

$$\tau_{\mathrm{DID}}$$

A statistical estimand
consistently estimated by $\hat{\tau}_{\mathrm{DID}}$

Identification: A Roadmap

$$\tau_{\text{DID}}$$

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

$$\tau_{\text{em}}$$

An associative estimand
describing effect heterogeneity

Identification: A Roadmap

$$\tau_{\text{DID}} \;\overset{\substack{\text{No anticipation \&} \\ \text{Parallel trends}}}{=\!=\!=} \;\tau_{\text{em}}$$

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

An associative estimand
describing effect heterogeneity

Identification: A Roadmap

No anticipation &
Parallel trends

τ_{DID} $=$ τ_{em} τ_{inter}

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

An associative estimand
describing effect heterogeneity

A causal estimand
targeting causal moderation

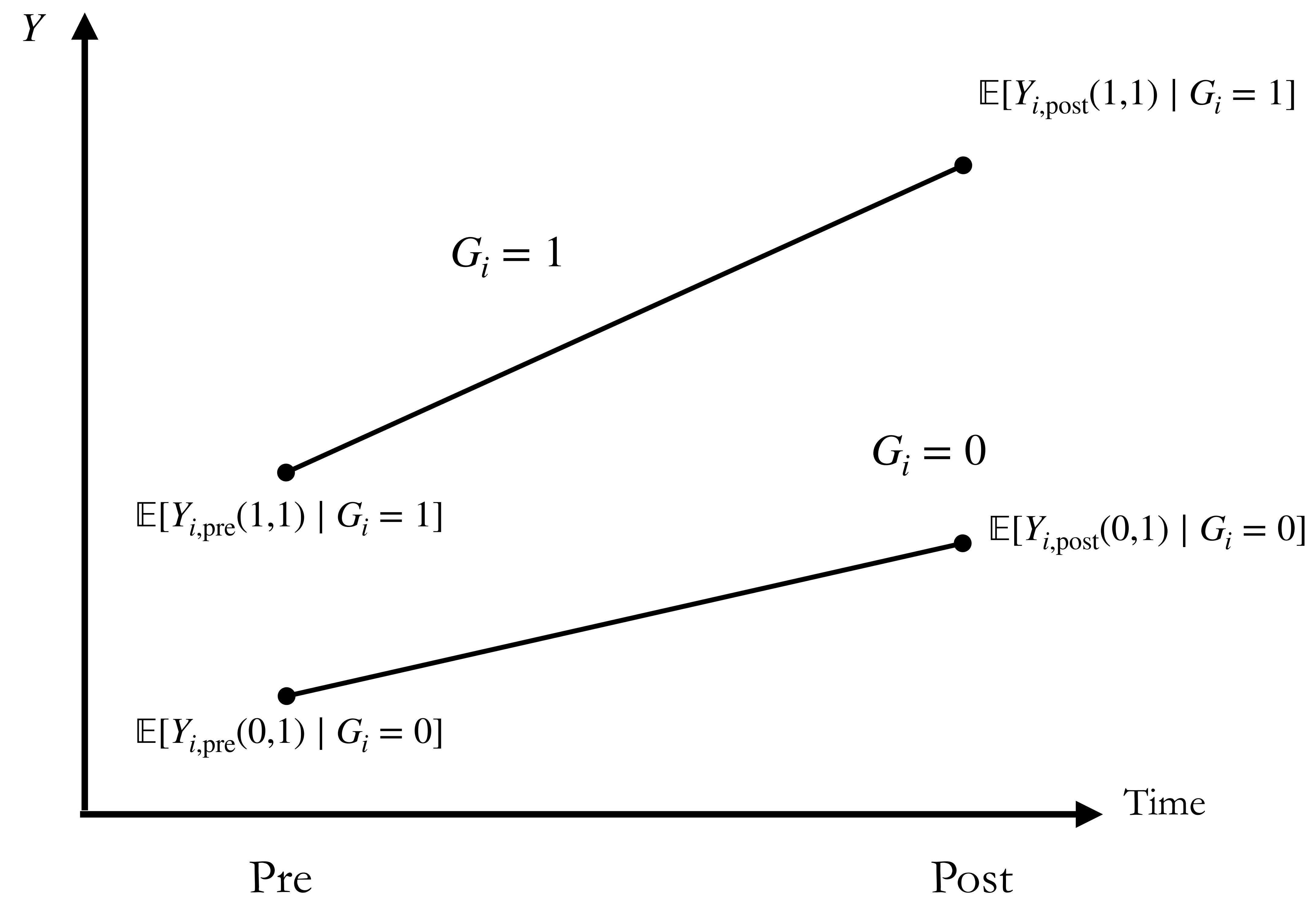
τ_{att}

A causal estimand
targeting the effect of Z
for G = 1

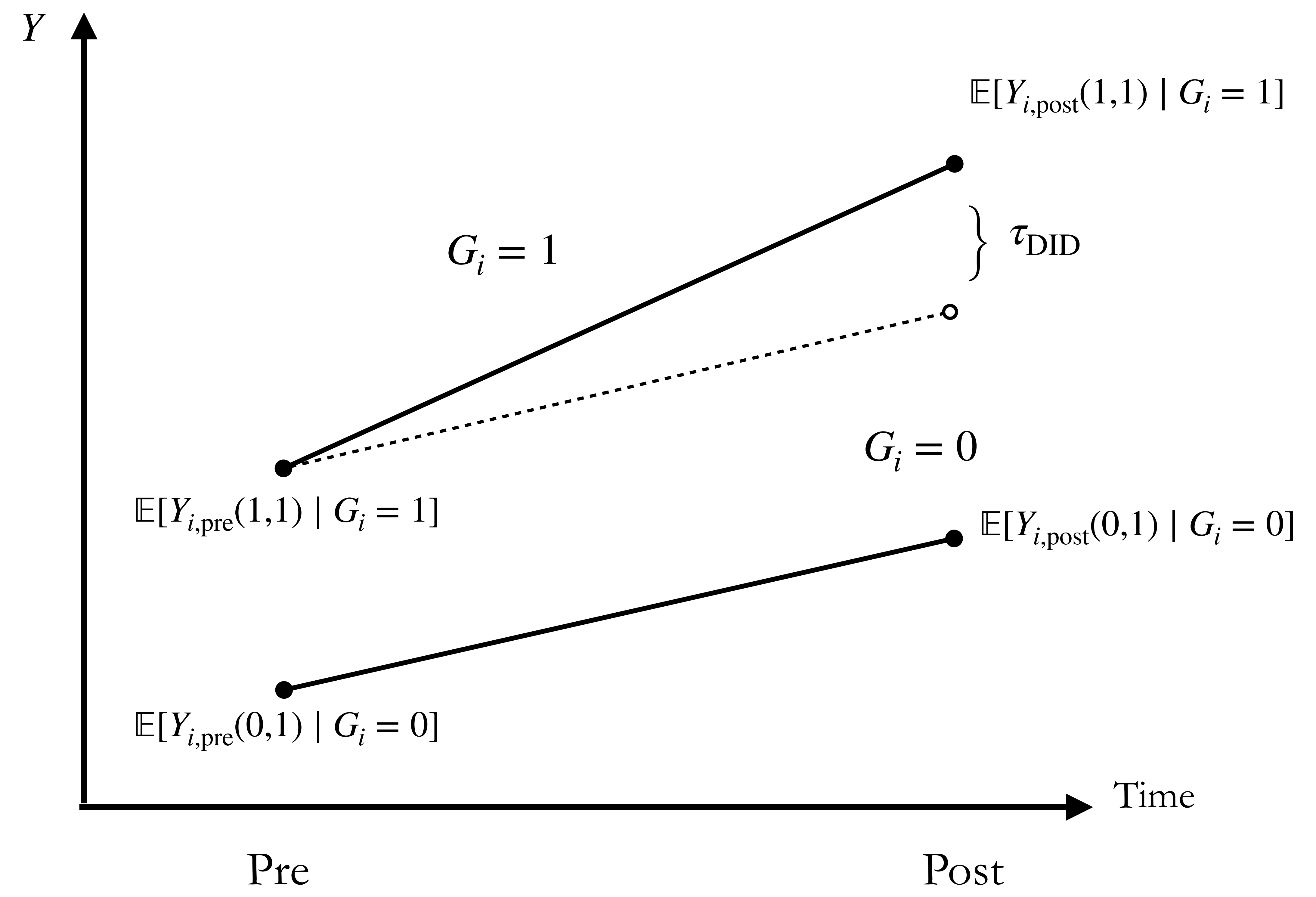
$\tau_{G|Z=1}$

A causal estimand
targeting the effect of G
when Z = 1

Identification under Canonical DID Assumptions



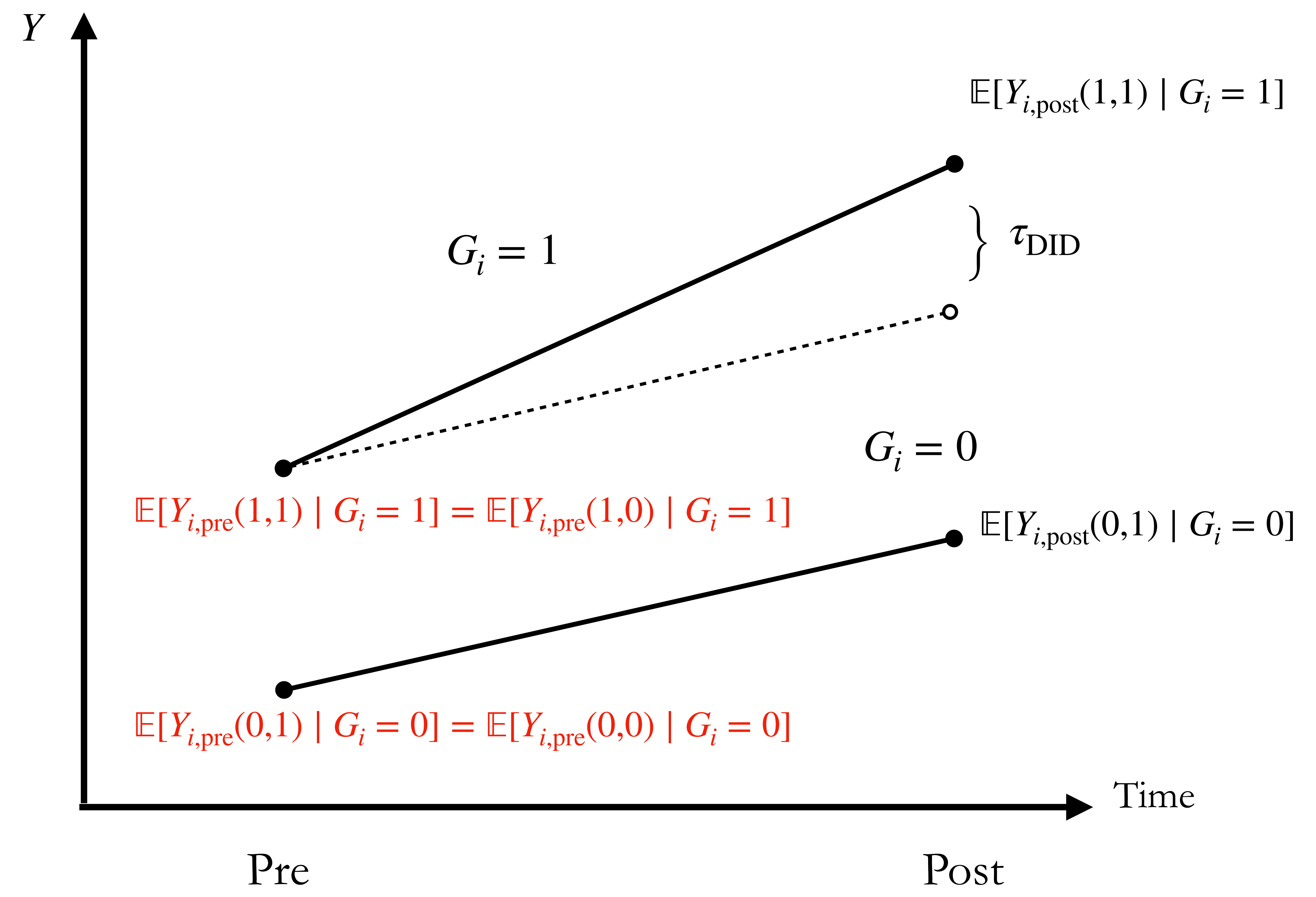
Identification under Canonical DID Assumptions



Identification under Canonical DID Assumptions

No Anticipation

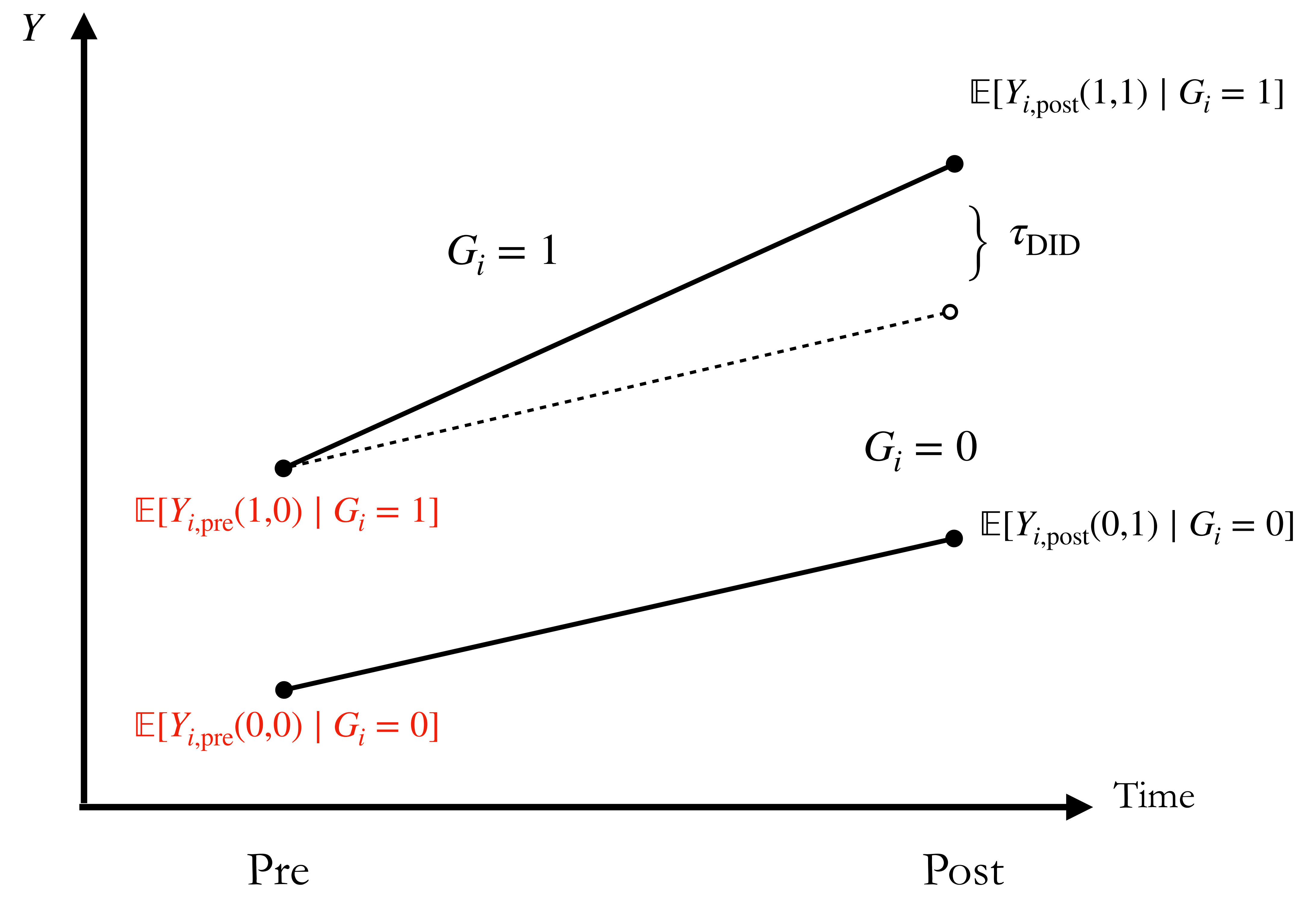
$Y_{i,\text{pre}}(g,0) = Y_{i,\text{pre}}(g,1)$ for all i and $g = 0,1$



Identification under Canonical DID Assumptions

No Anticipation

$Y_{i,\text{pre}}(g,0) = Y_{i,\text{pre}}(g,1)$ for all i and $g = 0,1$



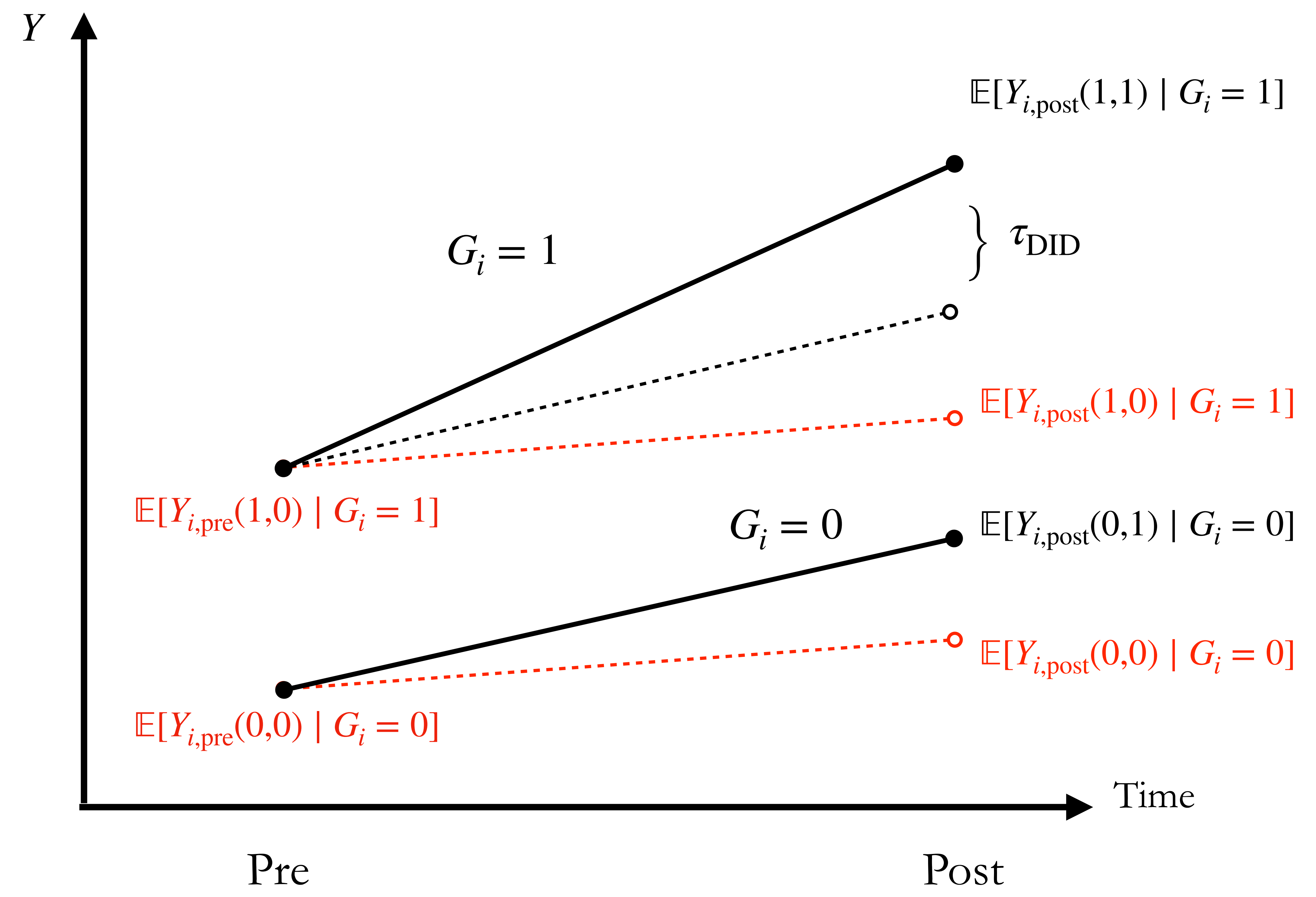
Identification under Canonical DID Assumptions

No Anticipation

$Y_{i,\text{pre}}(g,0) = Y_{i,\text{pre}}(g,1)$ for all i and $g = 0,1$

Parallel Trends

$\mathbb{E}[\Delta Y_i(1,0) \mid G_i = 1] = \mathbb{E}[\Delta Y_i(0,0) \mid G_i = 0]$



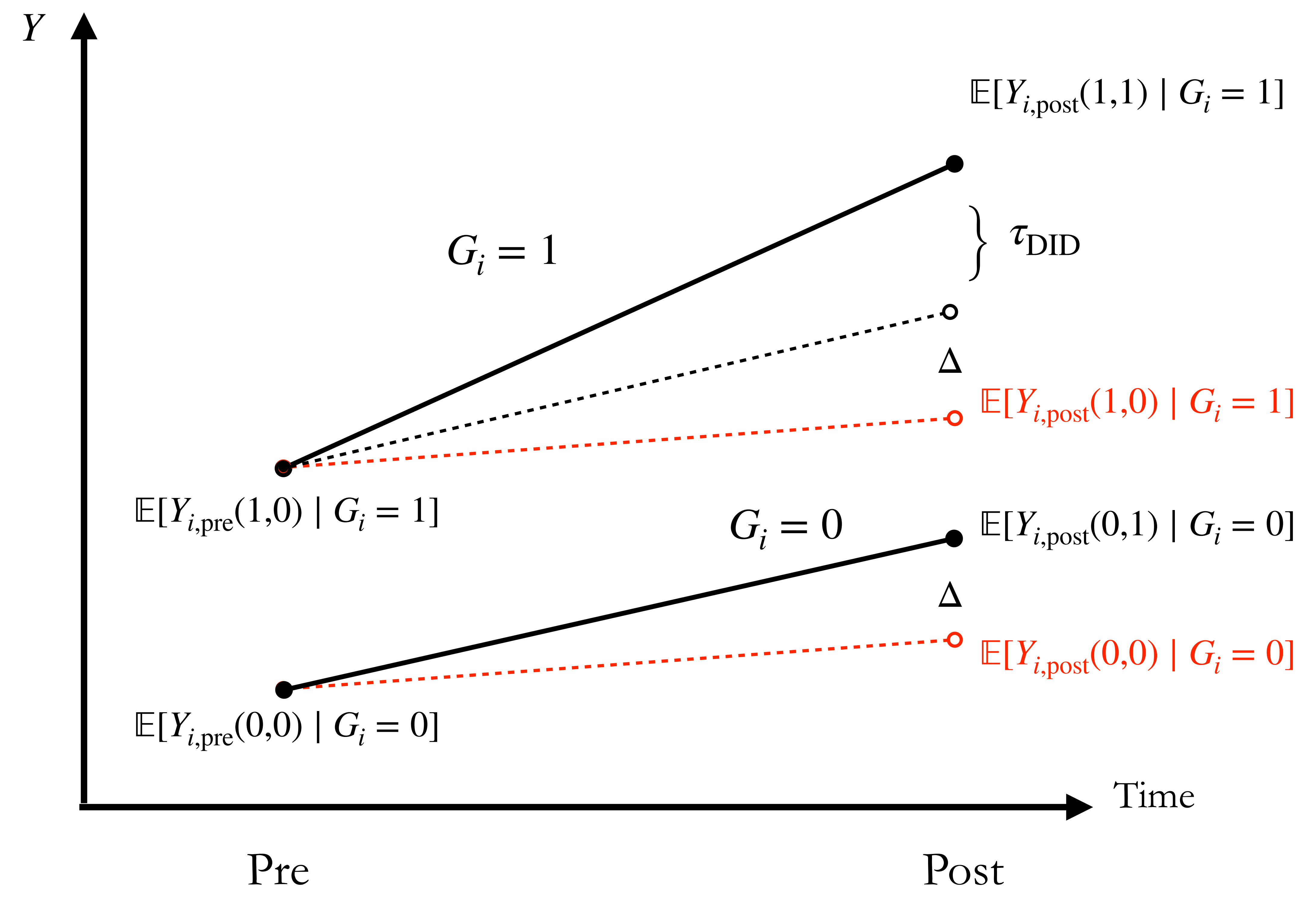
Identification under Canonical DID Assumptions

No Anticipation

$Y_{i,\text{pre}}(g,0) = Y_{i,\text{pre}}(g,1)$ for all i and $g = 0,1$

Parallel Trends

$\mathbb{E}[\Delta Y_i(1,0) \mid G_i = 1] = \mathbb{E}[\Delta Y_i(0,0) \mid G_i = 0]$



Identification under Canonical DID Assumptions

No Anticipation

$Y_{i,\text{pre}}(g,0) = Y_{i,\text{pre}}(g,1)$ for all i and $g = 0,1$

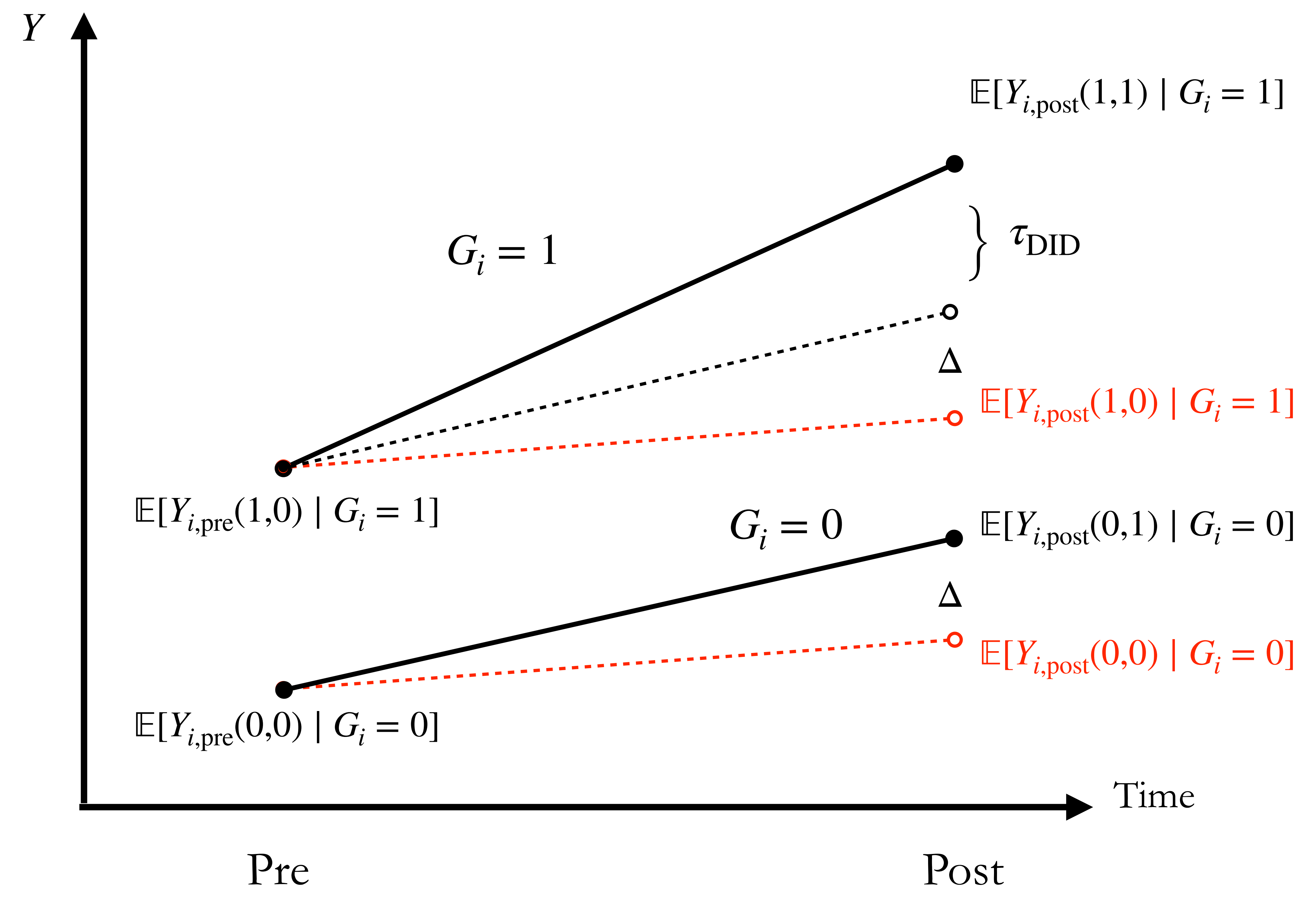
Parallel Trends

$\mathbb{E}[\Delta Y_i(1,0) \mid G_i = 1] = \mathbb{E}[\Delta Y_i(0,0) \mid G_i = 0]$

Proposition

Under no anticipation and parallel trends,

$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[\tau_{i,Z|G=1} \mid G_i = 1] - \mathbb{E}[\tau_{i,Z|G=0} \mid G_i = 0]$.



Identification Results

τ_{DID}

A statistical estimand consistently estimated by $\hat{\tau}_{\text{DID}}$

No anticipation &
Parallel trends

$$\tau_{\text{DID}} = \tau_{\text{em}}$$

τ_{em}

An associative estimand describing effect heterogeneity

?

τ_{inter}

A causal estimand targeting causal moderation

?

τ_{att}

A causal estimand targeting the effect of Z for G = 1

?

$\tau_{G|Z=1}$

A causal estimand targeting the effect of G when Z = 1

Identification Results

No anticipation &
Parallel trends

τ_{DID} $=$ τ_{em} ? τ_{inter}

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

An associative estimand
describing effect heterogeneity

A causal estimand
targeting causal moderation

?

?

τ_{att}

A causal estimand
targeting the effect of Z
for G = 1

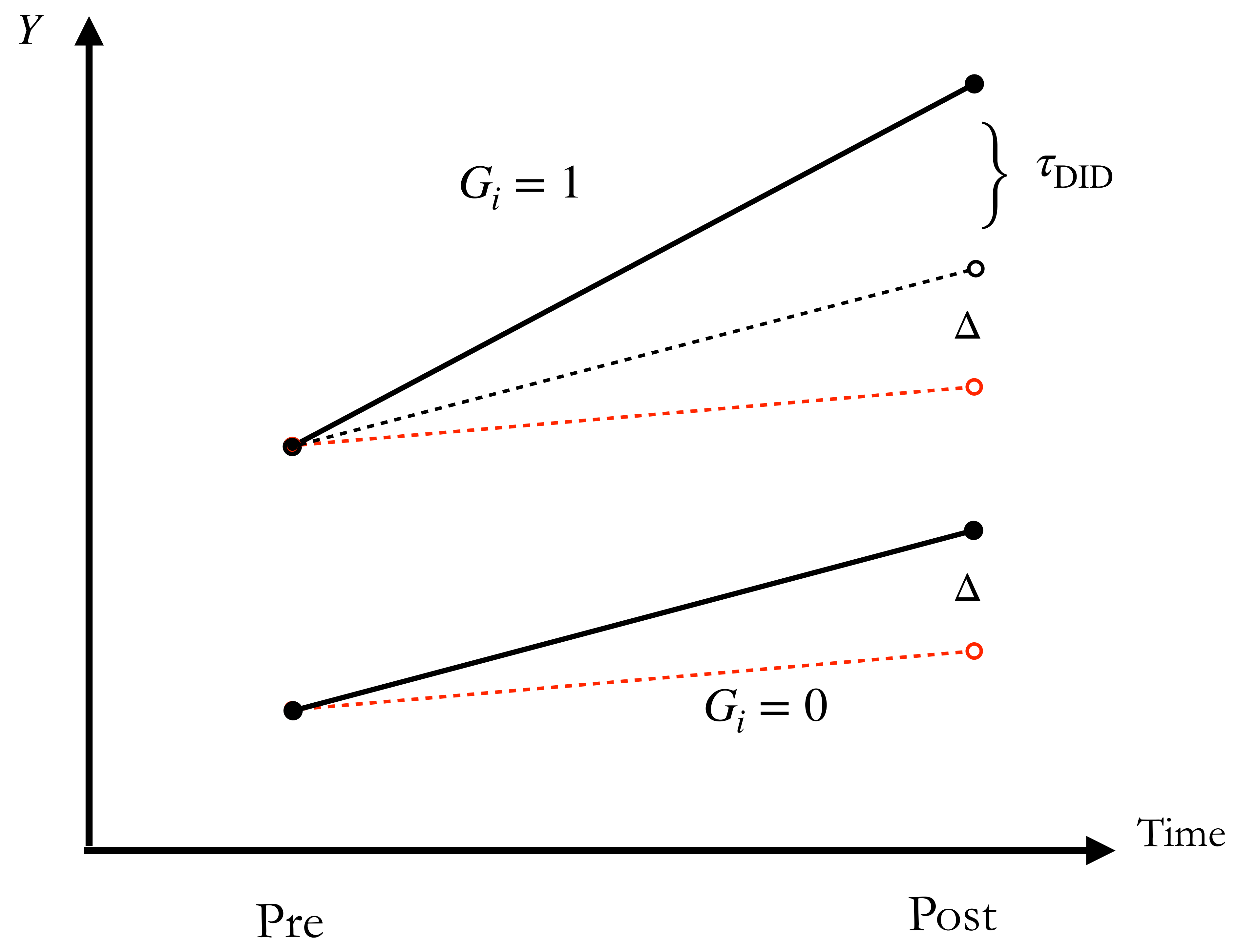
$\tau_{G|Z=1}$

A causal estimand
targeting the effect of G
when Z = 1

Reconciling Canonical and Factorial DID

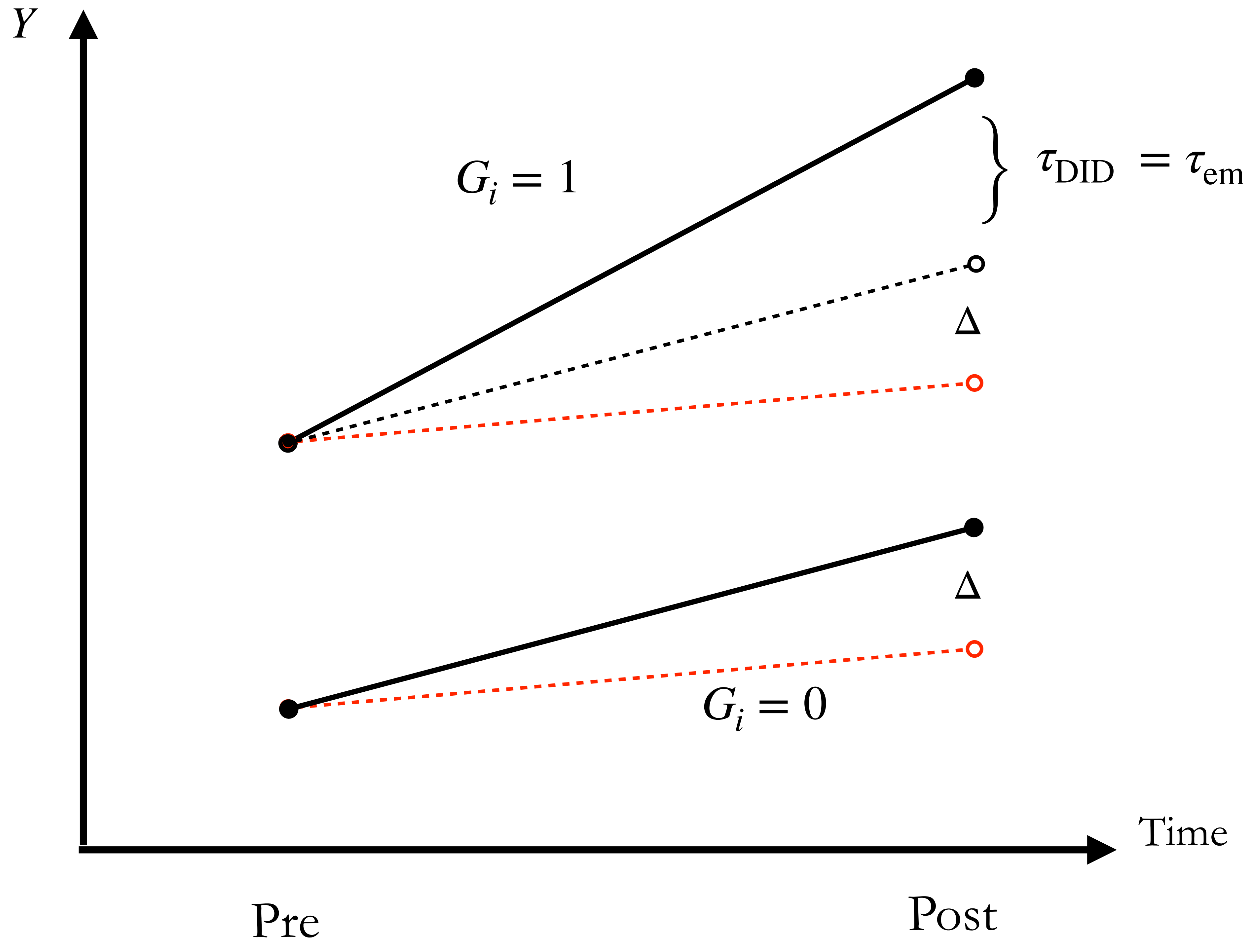
Reconciling Canonical and Factorial DID

Factorial DID



Reconciling Canonical and Factorial DID

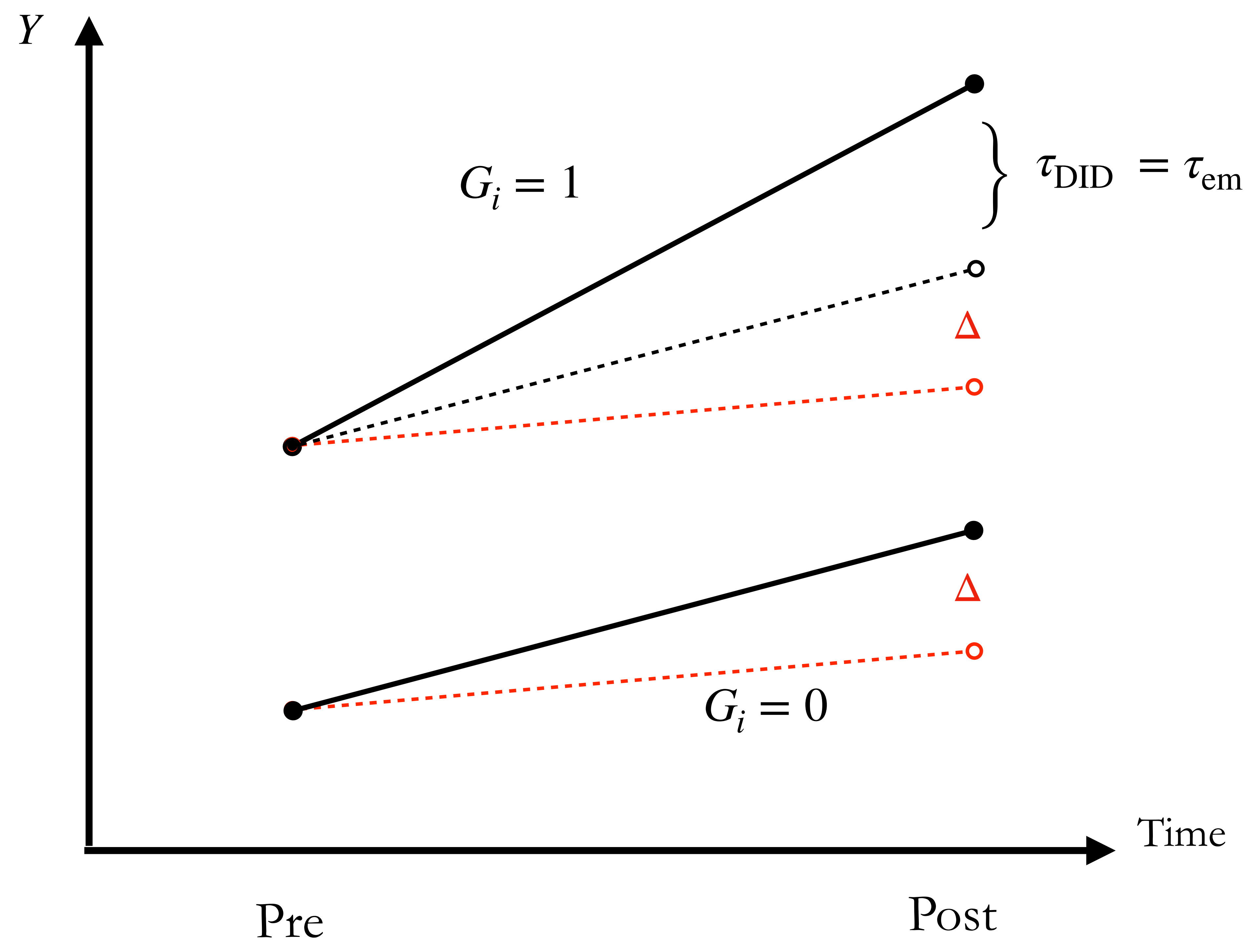
Factorial DID



$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1]$$
$$- \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

Reconciling Canonical and Factorial DID

Factorial DID



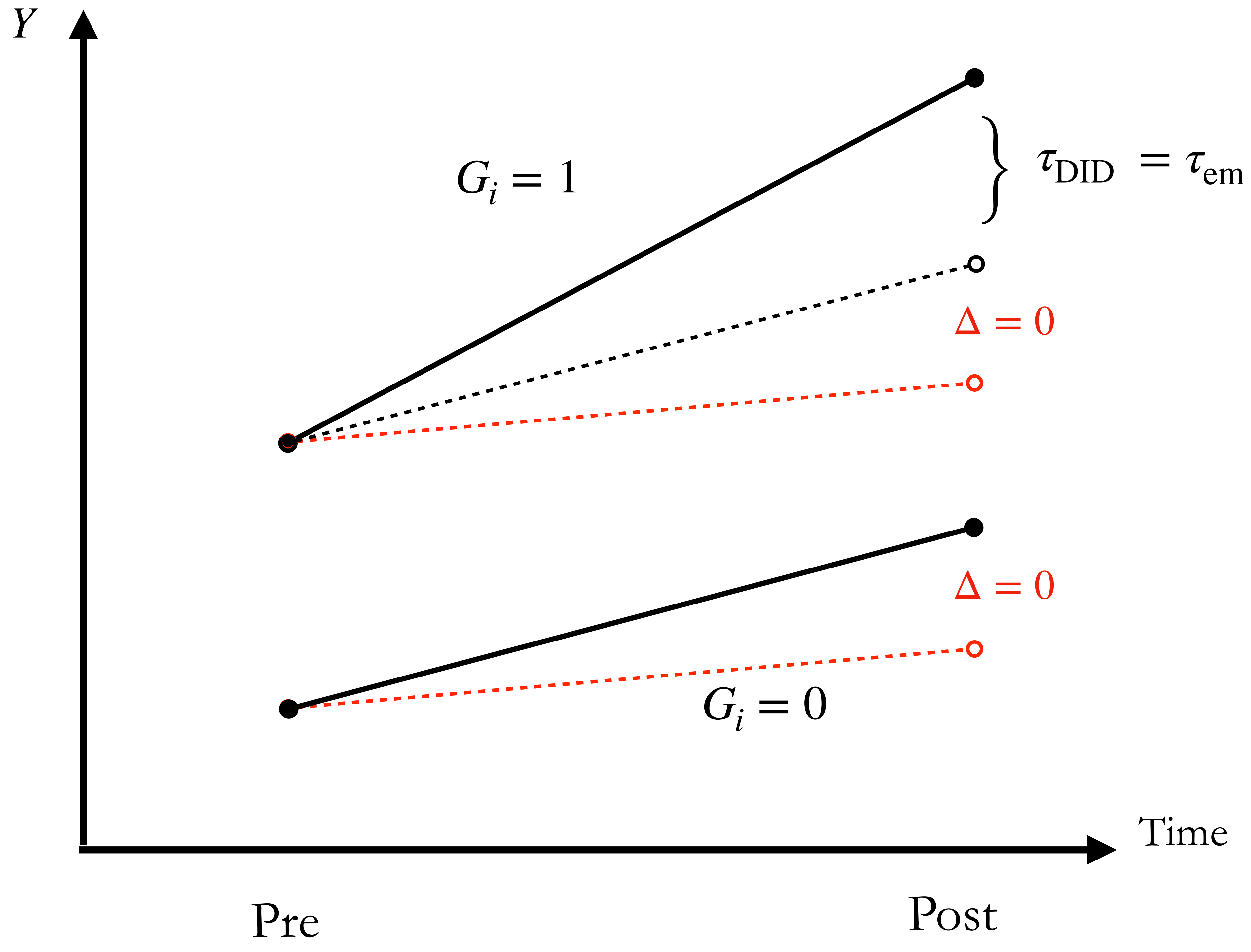
$$\tau_{\mathrm{DID}} = \tau_{\mathrm{em}} = \mathbb{E}[Y_{i,\mathrm{post}}(1,1) - Y_{i,\mathrm{post}}(1,0) \mid G_i = 1]$$
$$- \mathbb{E}[Y_{i,\mathrm{post}}(0,1) - Y_{i,\mathrm{post}}(0,0) \mid G_i = 0]$$

Exclusion Restriction on Z

$Y_{i,\mathrm{post}}(0,1) = Y_{i,\mathrm{post}}(0,0)$ for all units with $G_i = 0$

Reconciling Canonical and Factorial DID

Factorial DID



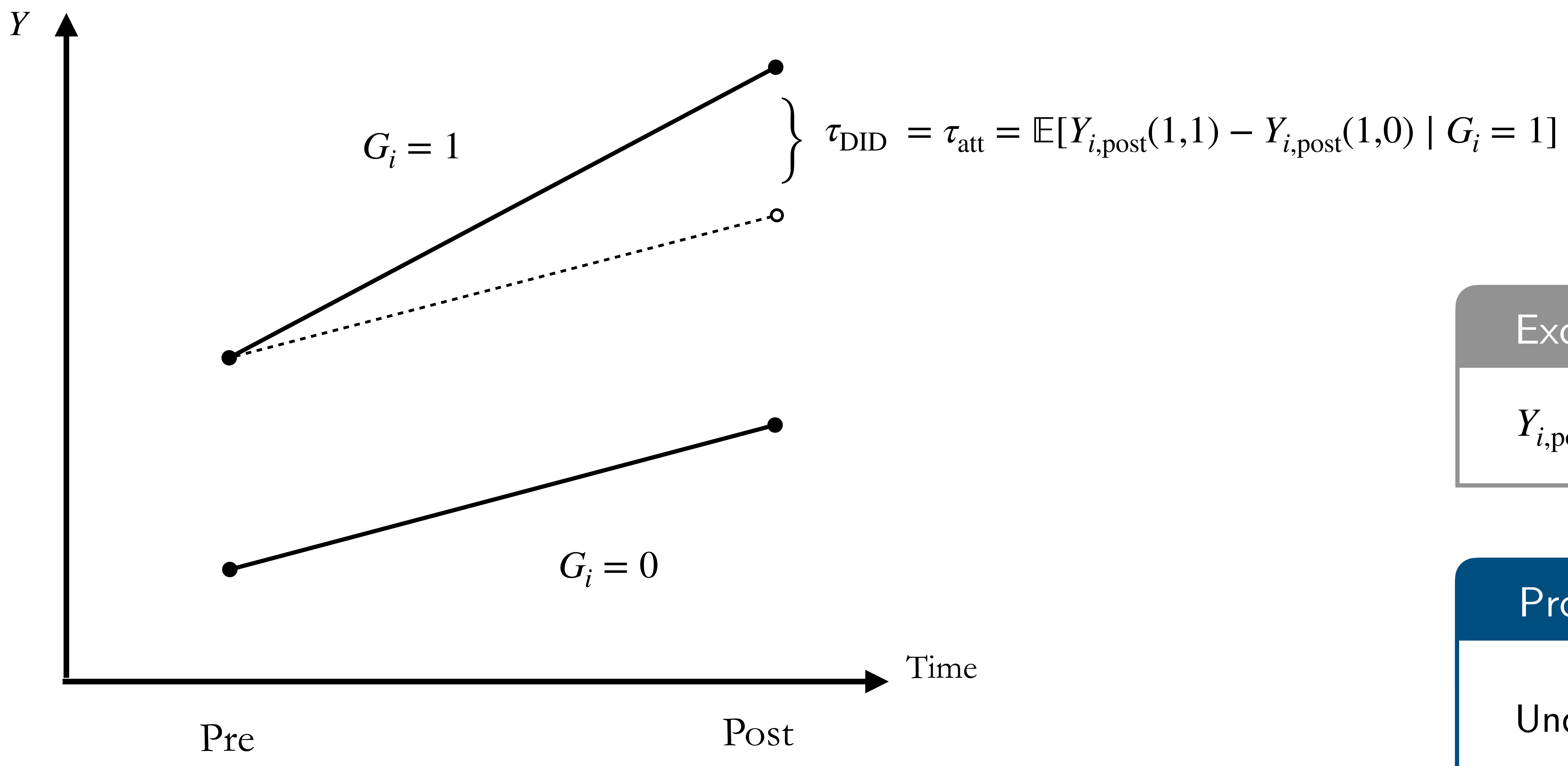
$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1]$$
$$-\mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

Exclusion Restriction on Z

$Y_{i,\text{post}}(0,1) = Y_{i,\text{post}}(0,0)$ for all units with $G_i = 0$

Reconciling Canonical and Factorial DID

Factorial DID → Canonical DID



Exclusion Restriction on Z

$Y_{i,\text{post}}(0,1) = Y_{i,\text{post}}(0,0)$ for all units with $G_i = 0$

Proposition

Under no anticipation, parallel trends, and the exclusion restriction, $\tau_{\text{DID}} = \tau_{\text{att}}$.

Identification Results

τ_{DID}

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

No anticipation &
Parallel trends

$$=$$

τ_{em}

An associative estimand
describing effect heterogeneity

τ_{inter}

A causal estimand
targeting causal moderation

τ_{att}

A causal estimand
targeting the effect of Z
for G = 1

$\tau_{G|Z=1}$

A causal estimand
targeting the effect of G
when Z = 1

Identification Results

No anticipation &
Parallel trends

τ_{DID} $=$ τ_{em} $\quad$ τ_{inter}

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

An associative estimand
describing effect heterogeneity

A causal estimand
targeting causal moderation

Exclusion restriction on Z
for G = 0

$\|$

τ_{att}

A causal estimand
targeting the effect of Z
for G = 1

$\tau_{G|Z=1}$

A causal estimand
targeting the effect of G
when Z = 1

Identification Results

No anticipation &
Parallel trends

τ_{DID} $=$ τ_{em} **?** τ_{inter}

A statistical estimand
consistently estimated by $\hat{\tau}_{\text{DID}}$

An associative estimand
describing effect heterogeneity

A causal estimand
targeting causal moderation

Exclusion restriction on Z
for G = 0

$\| $

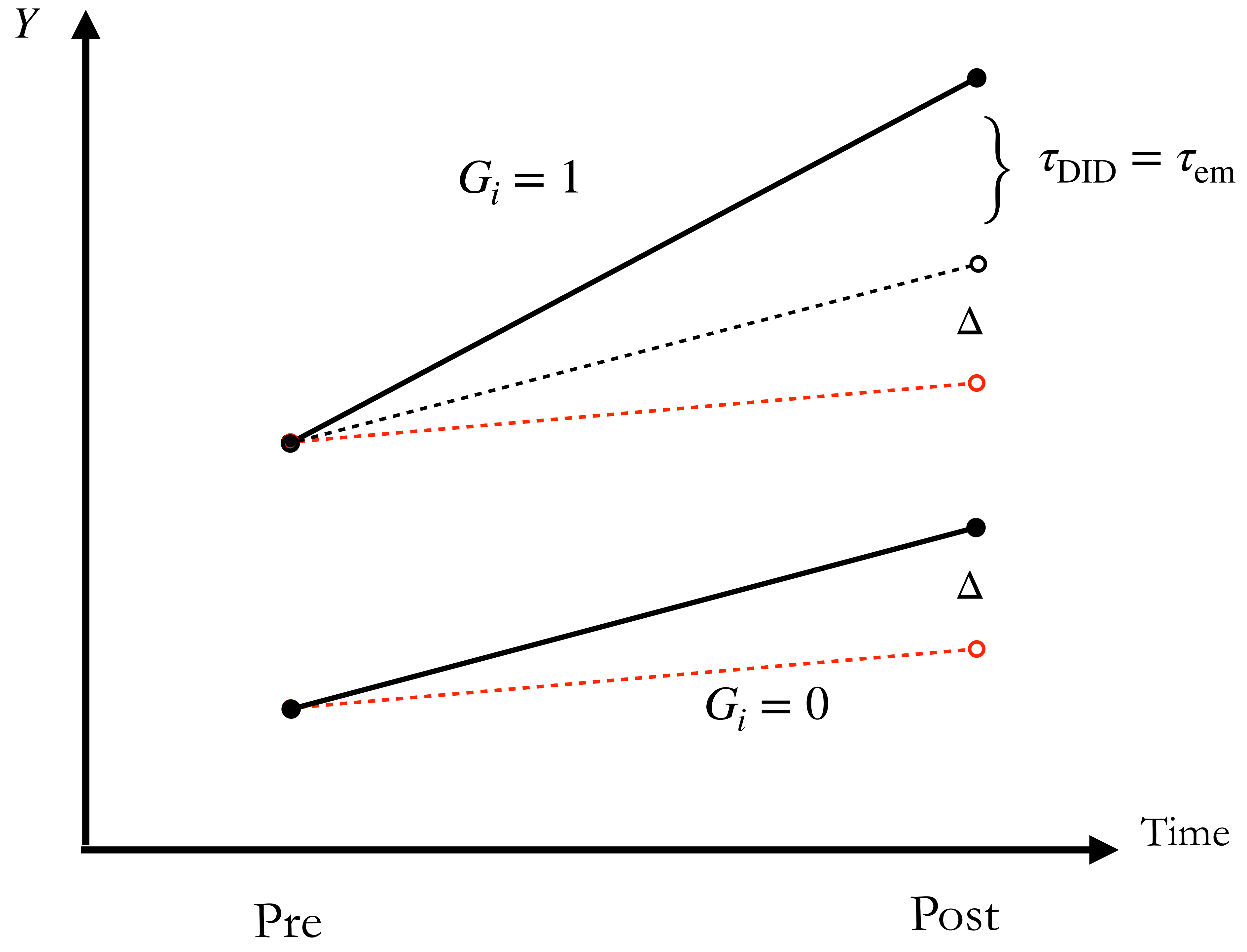
τ_{att}

A causal estimand
targeting the effect of Z
for G = 1

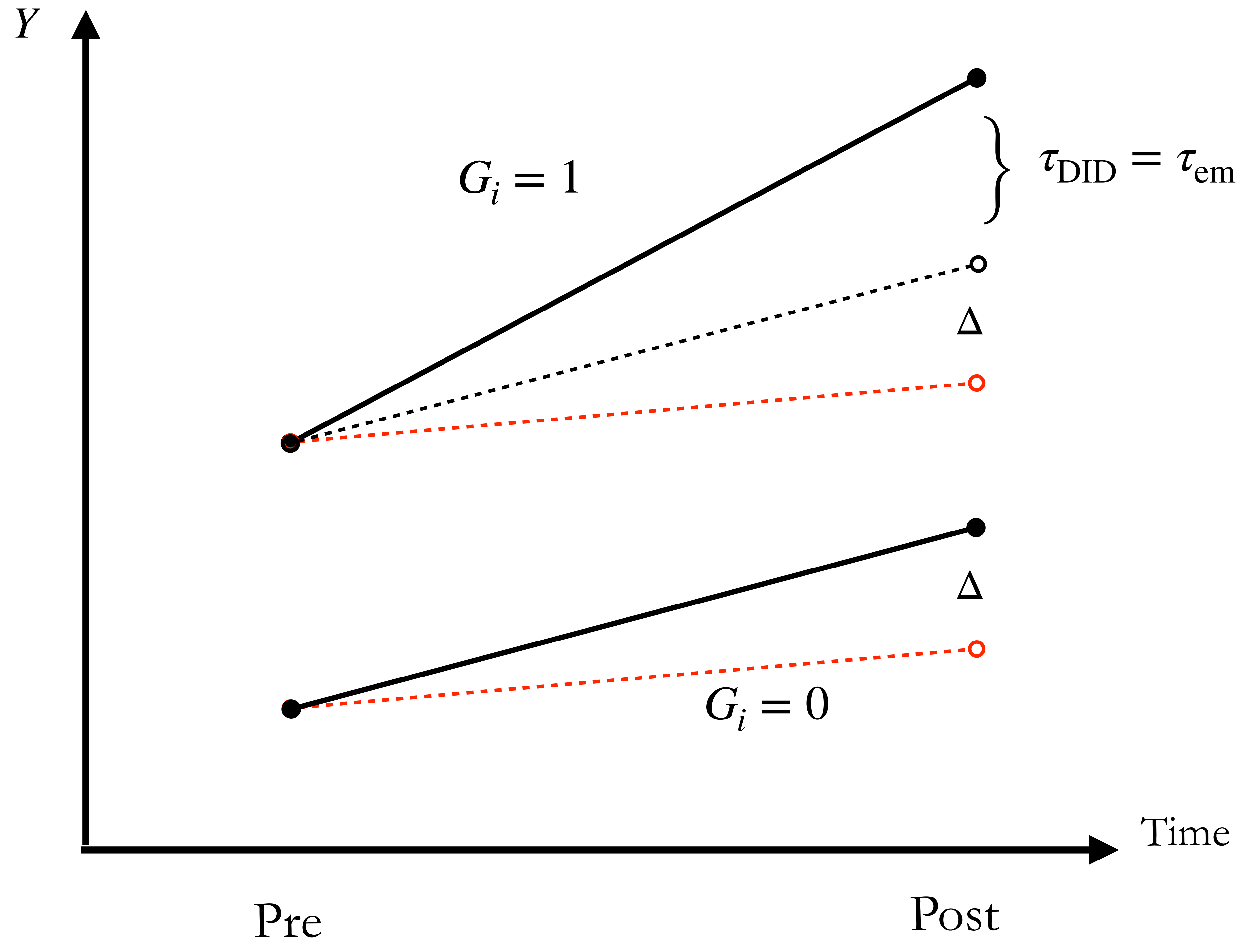
$\tau_{G|Z=1}$

A causal estimand
targeting the effect of G
when Z = 1

Why τ_{DID} May Not Have a Causal Interpretation under Parallel Trends?

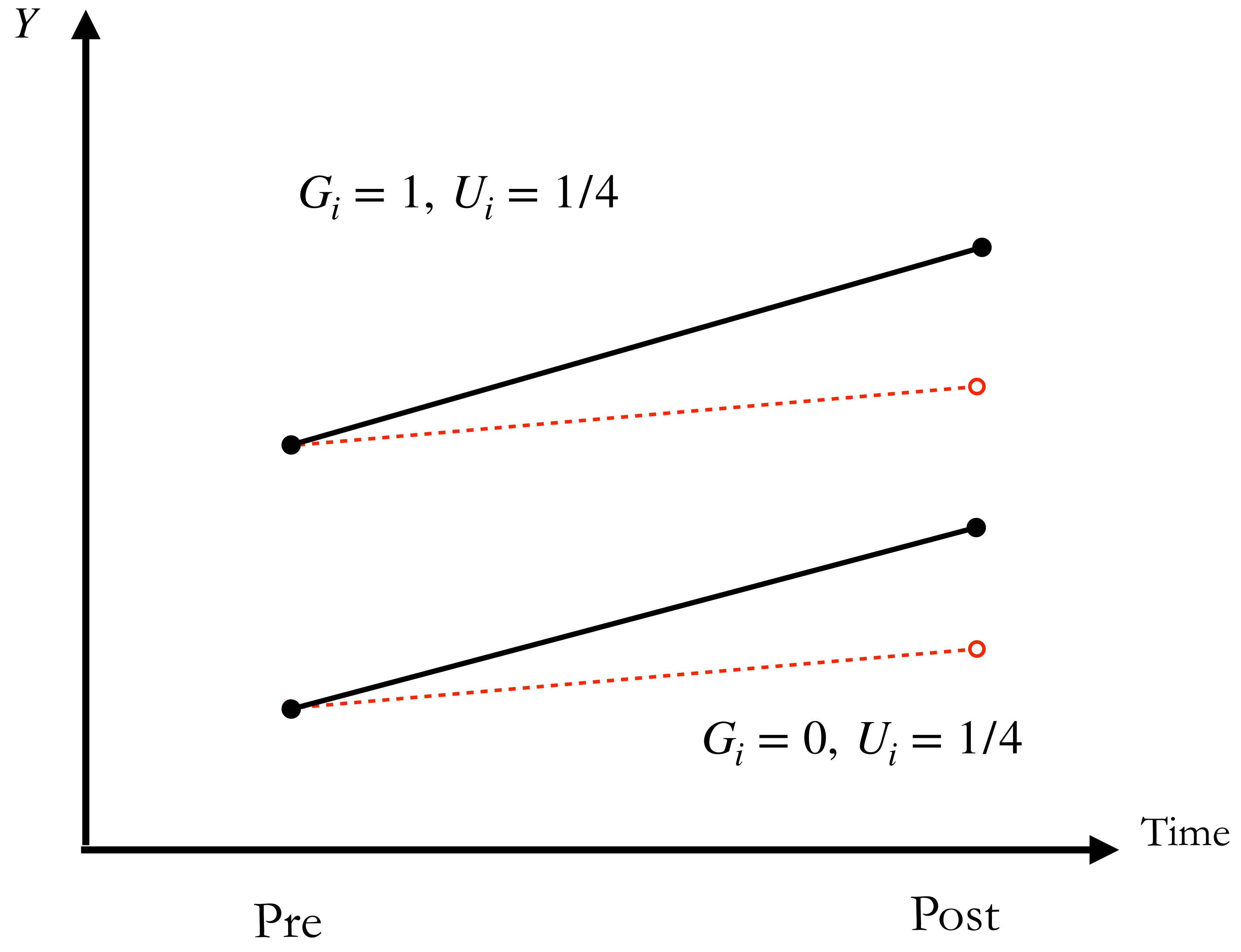


Why τ_{DID} May Not Have a Causal Interpretation under Parallel Trends?



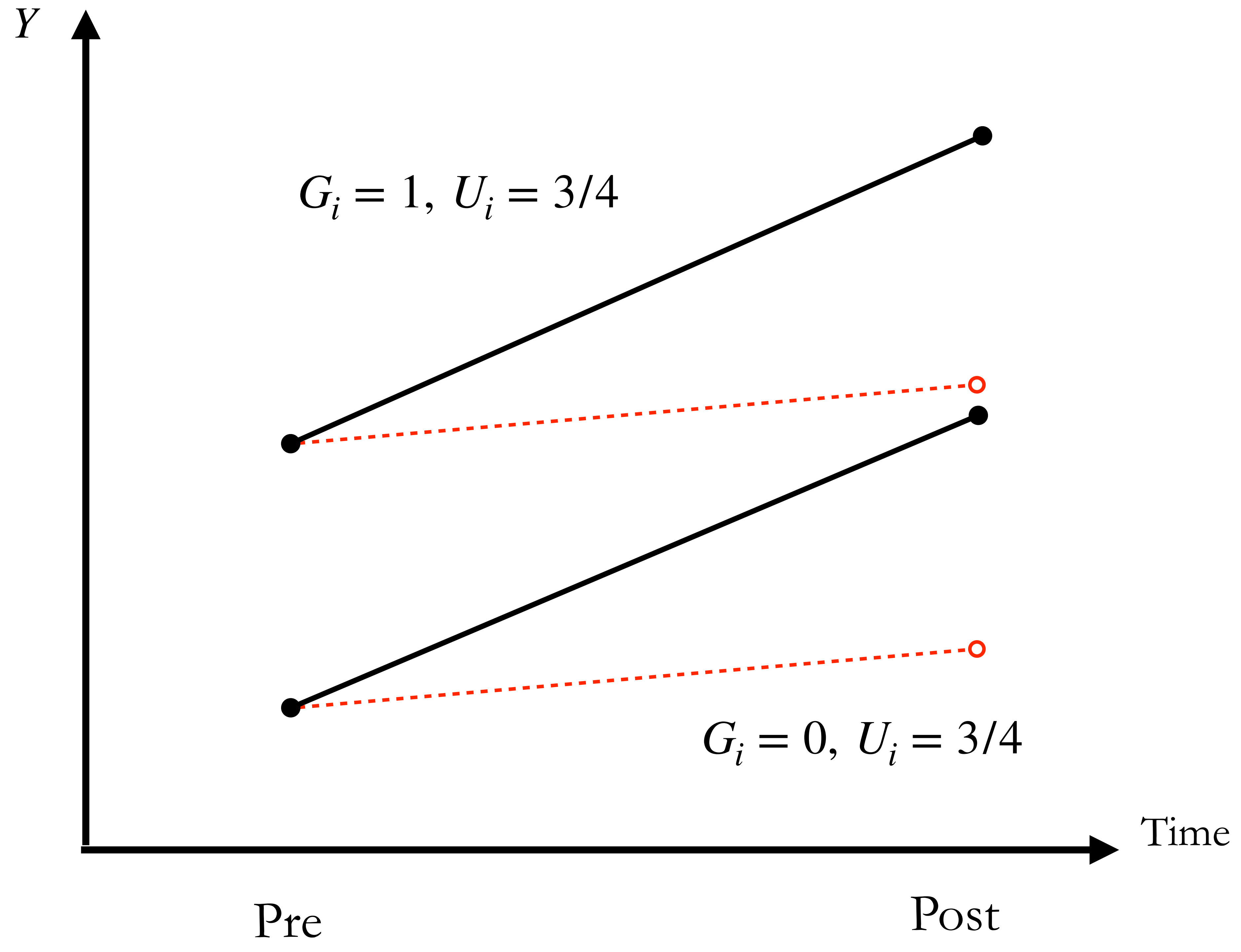
- Imagine an unobservable U that determines how units respond to the event

Why τ_{DID} May Not Have a Causal Interpretation under Parallel Trends?



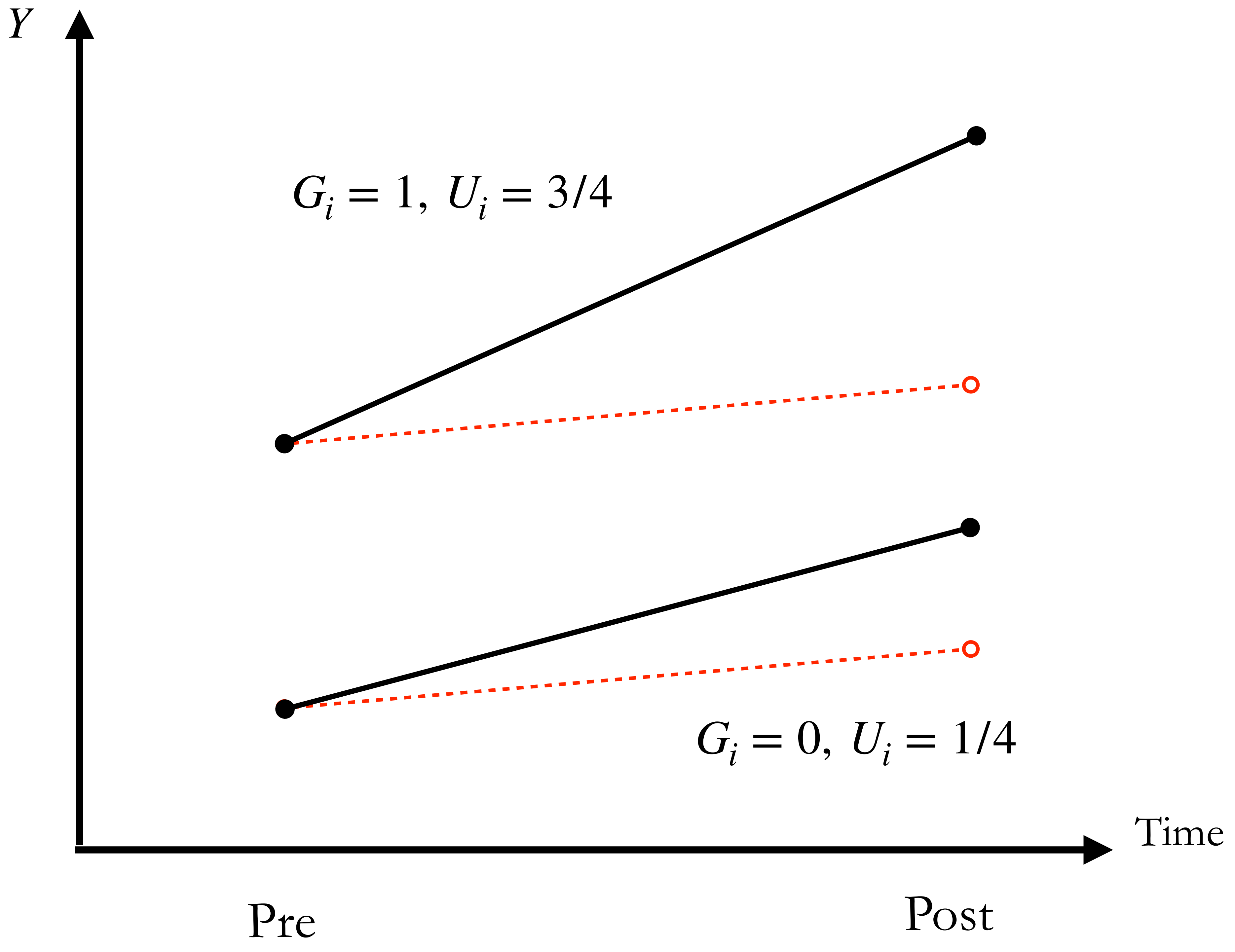
- Imagine an unobservable U that determines how units respond to the event

Why τ_{DID} May Not Have a Causal Interpretation under Parallel Trends?



- Imagine an unobservable U that determines how units respond to the event

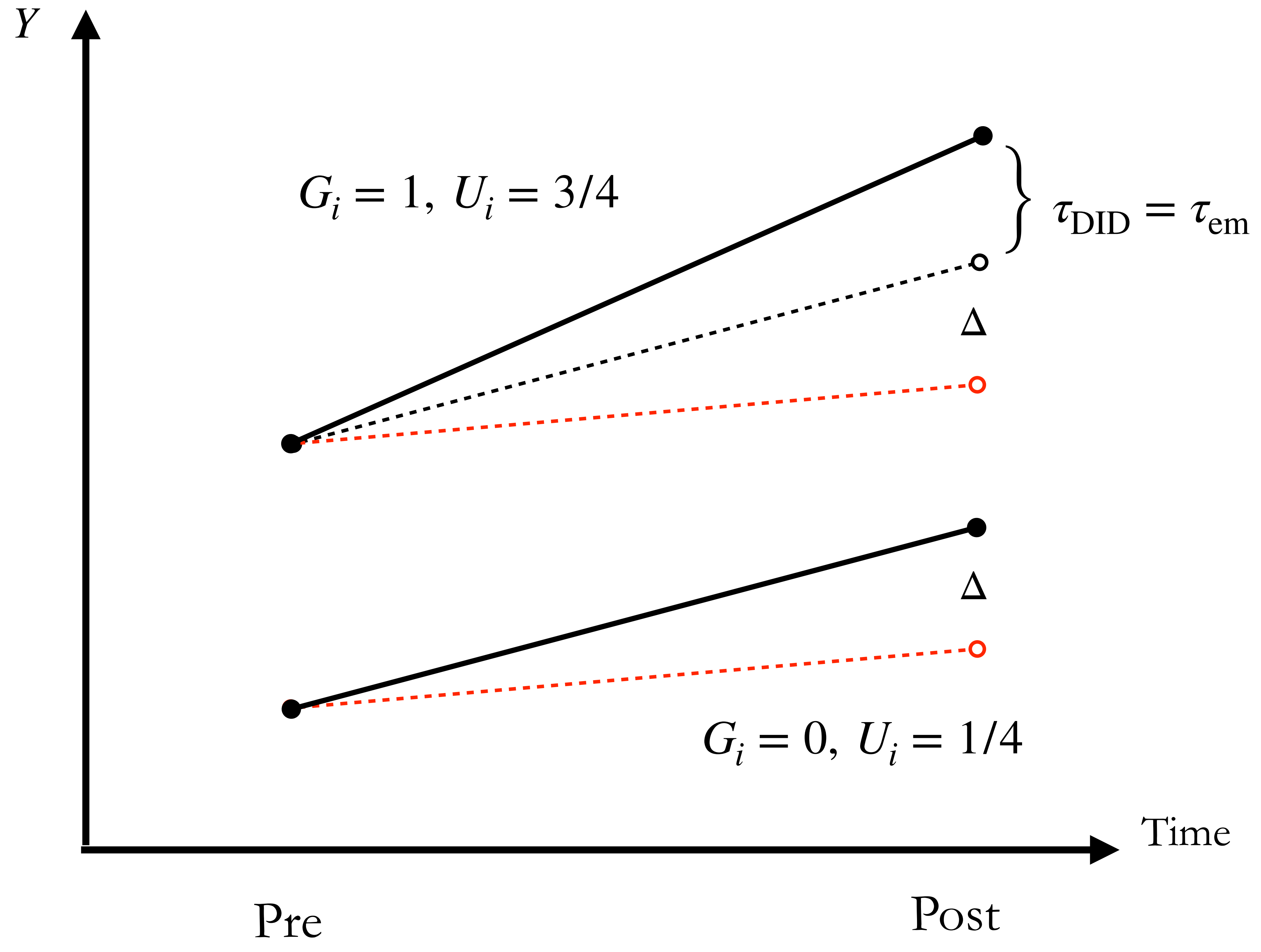
Why τ_{DID} May Not Identify Causal Moderation under Parallel Trends?



- Imagine an unobservable U that determines how units respond to the event

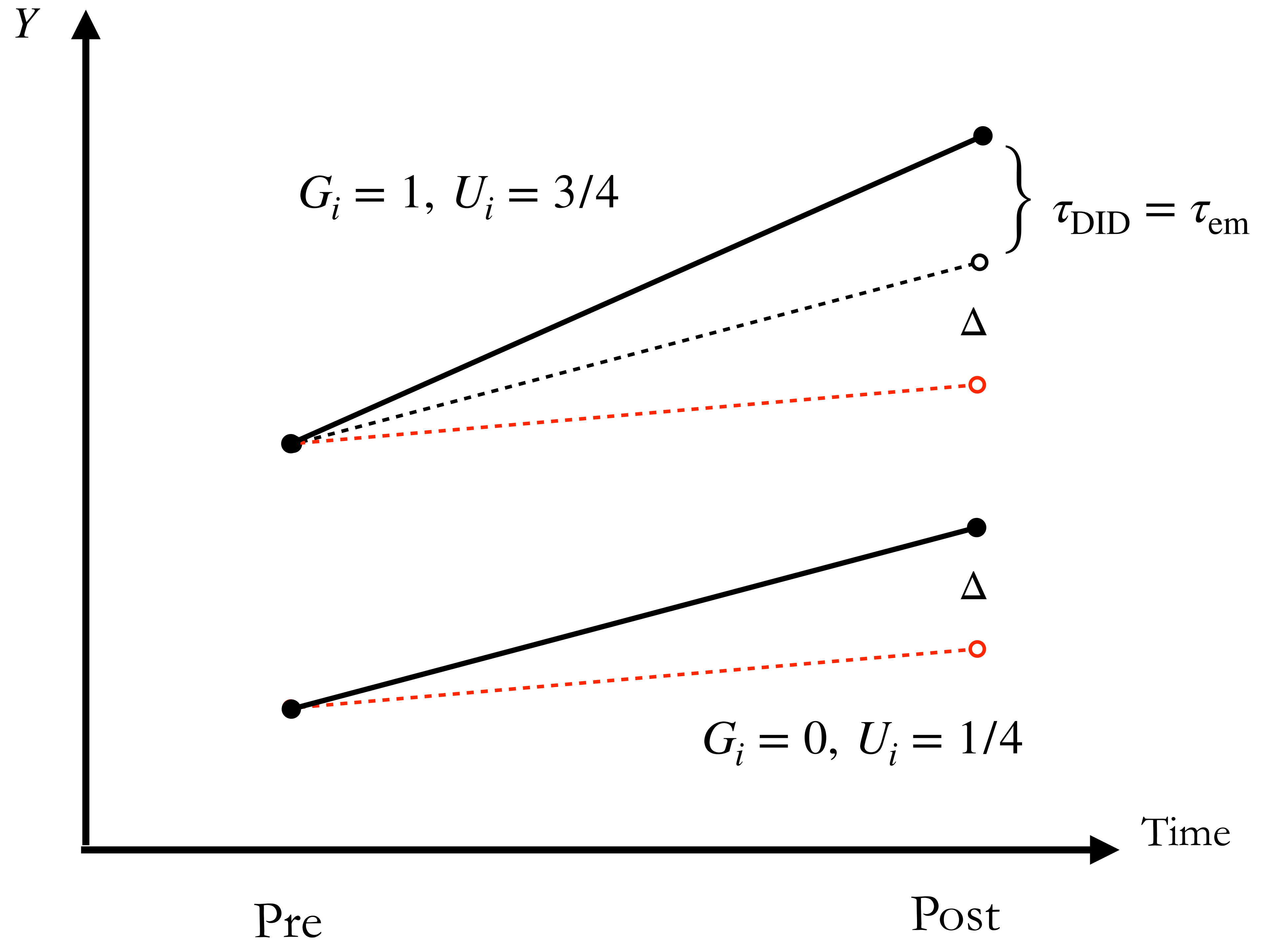
- U may be correlated with G

Why τ_{DID} May Not Identify Causal Moderation under Parallel Trends?



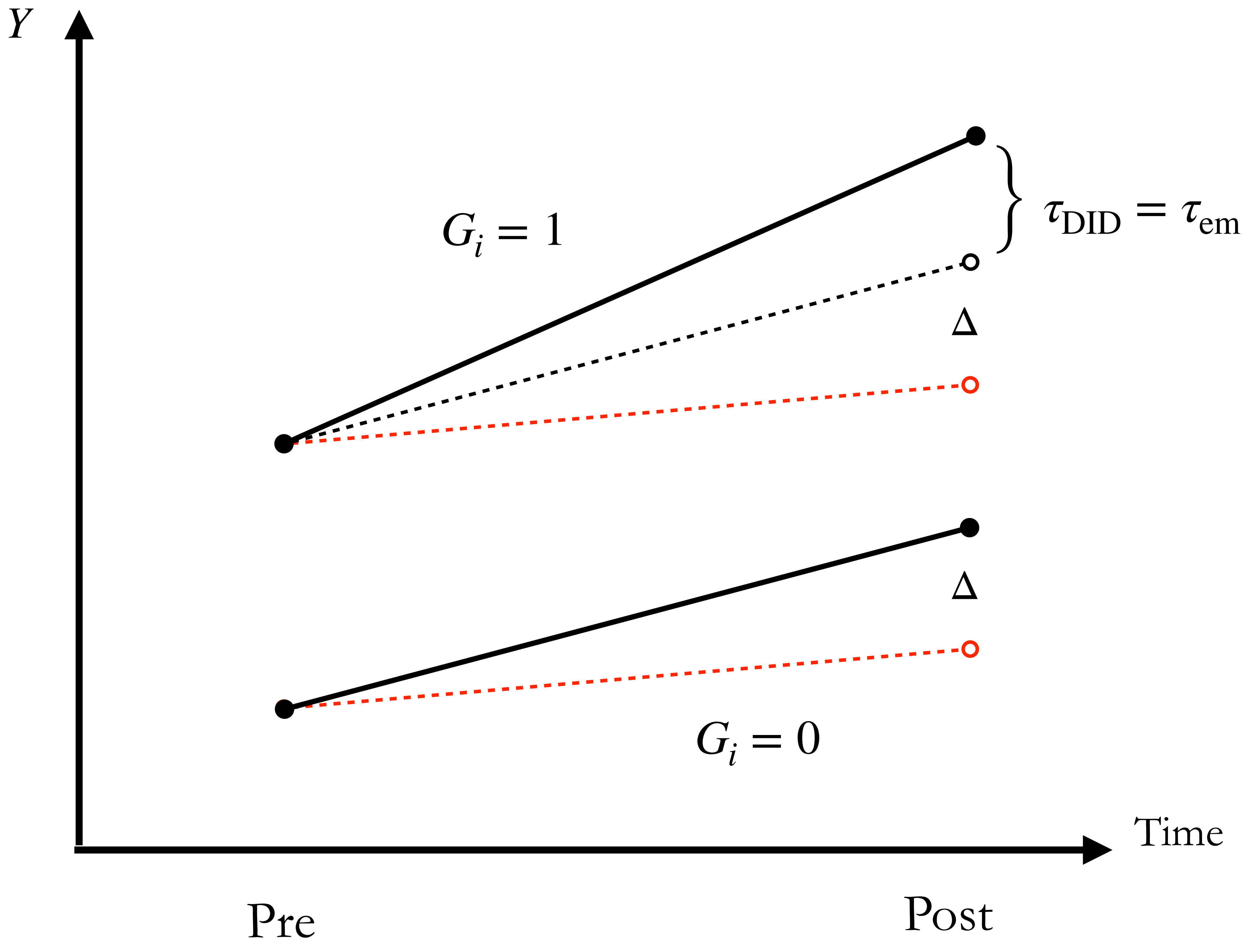
- Imagine an unobservable U that determines how units respond to the event
- U may be correlated with G

Why τ_{DID} May Not Identify Causal Moderation under Parallel Trends?



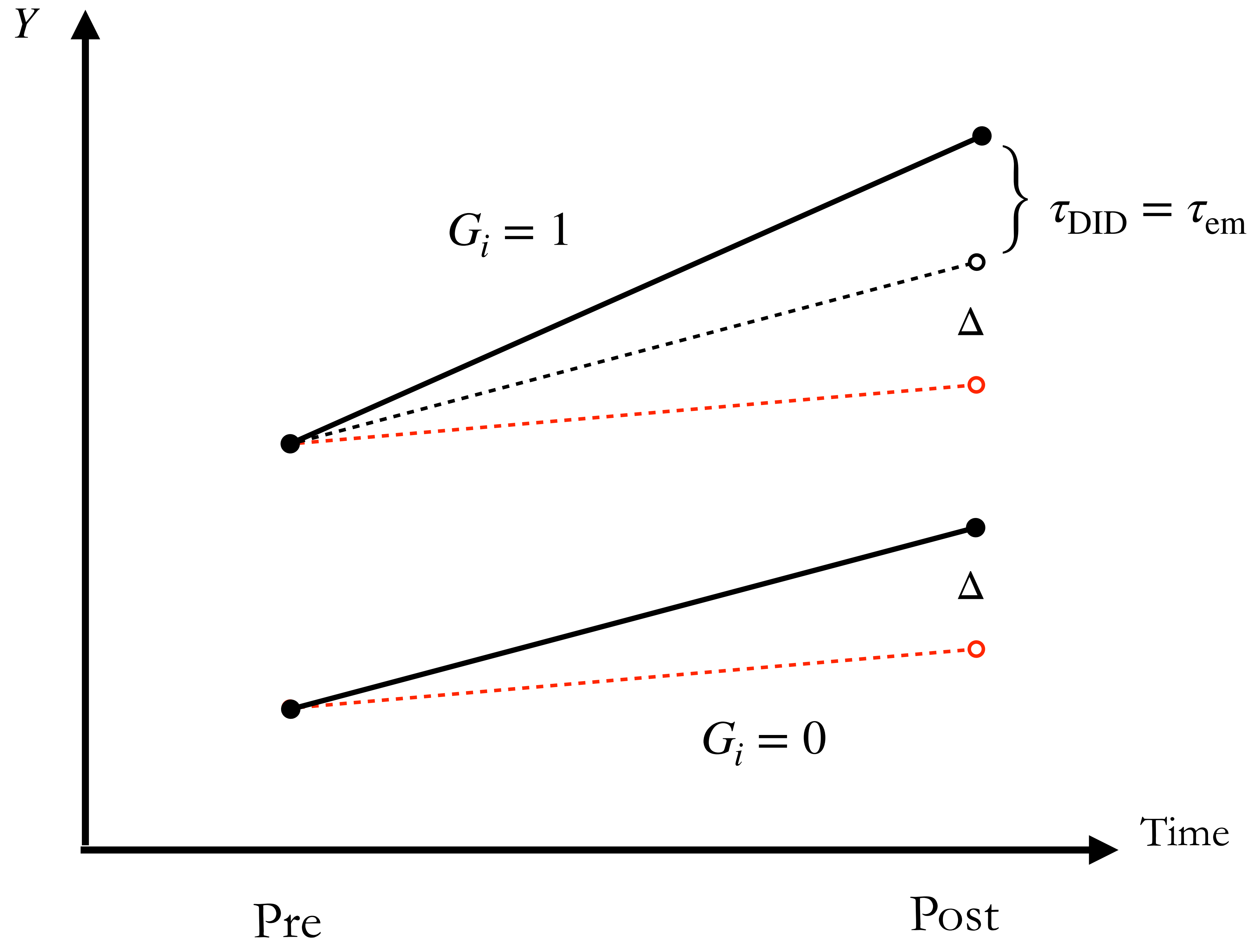
- Imagine an unobservable U that determines how units respond to the event
- U may be correlated with G
- Therefore, τ_{DID} cannot be interpreted as the causal moderation of G

Identifying Causal Interaction (Causal Moderation)



$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

Identifying Causal Interaction (Causal Moderation)

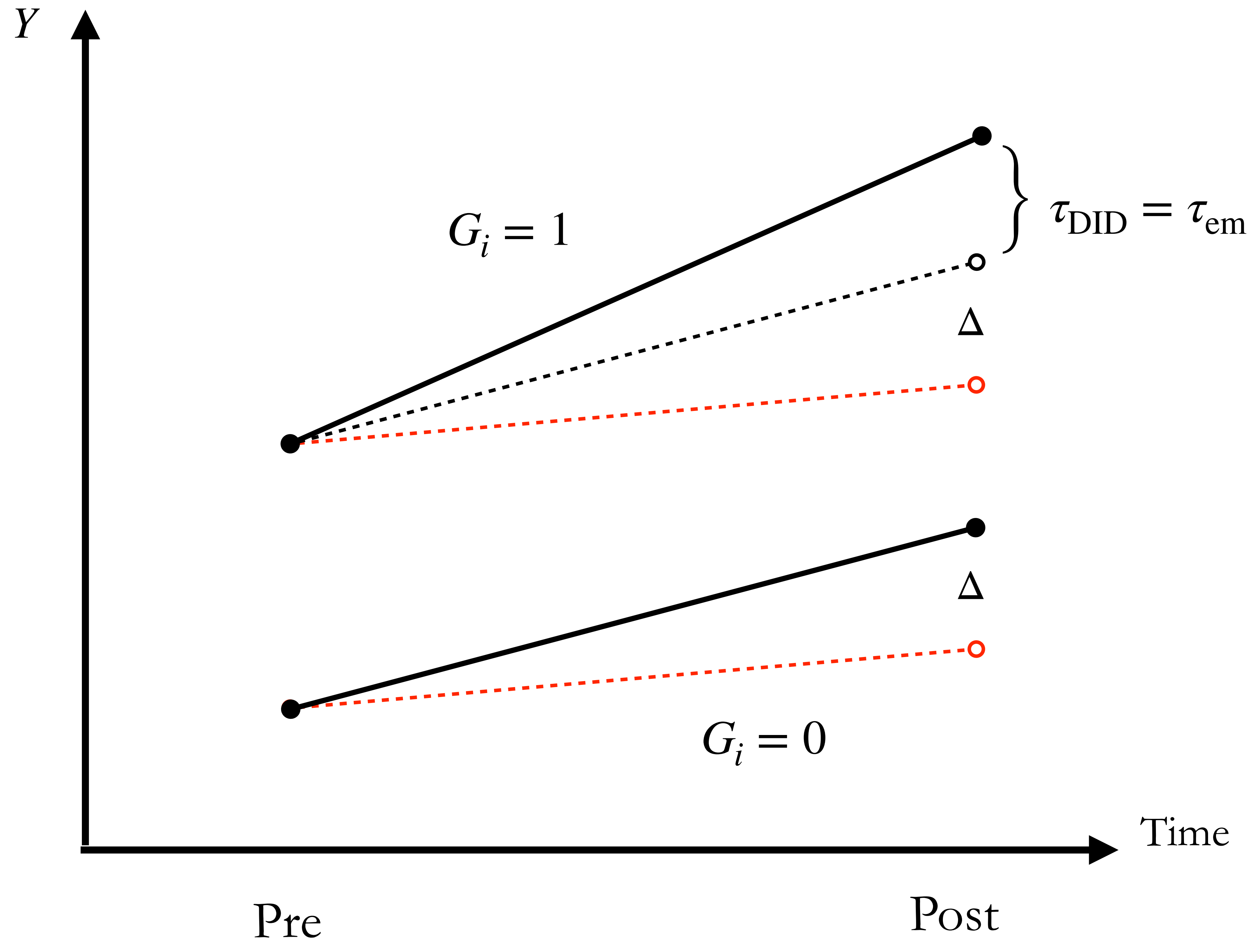


$$\tau_{\mathrm{DID}} = \tau_{\mathrm{em}} = \mathbb{E}[Y_{i,\mathrm{post}}(1,1) - Y_{i,\mathrm{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\mathrm{post}}(0,1) - Y_{i,\mathrm{post}}(0,0) \mid G_i = 0]$$

$$\overset{?}{=}$$

$$\tau_{\mathrm{inter}} = \mathbb{E}[Y_{i,\mathrm{post}}(1,1) - Y_{i,\mathrm{post}}(1,0)] - \mathbb{E}[Y_{i,\mathrm{post}}(0,1) - Y_{i,\mathrm{post}}(0,0)]$$

Identifying Causal Interaction (Causal Moderation)



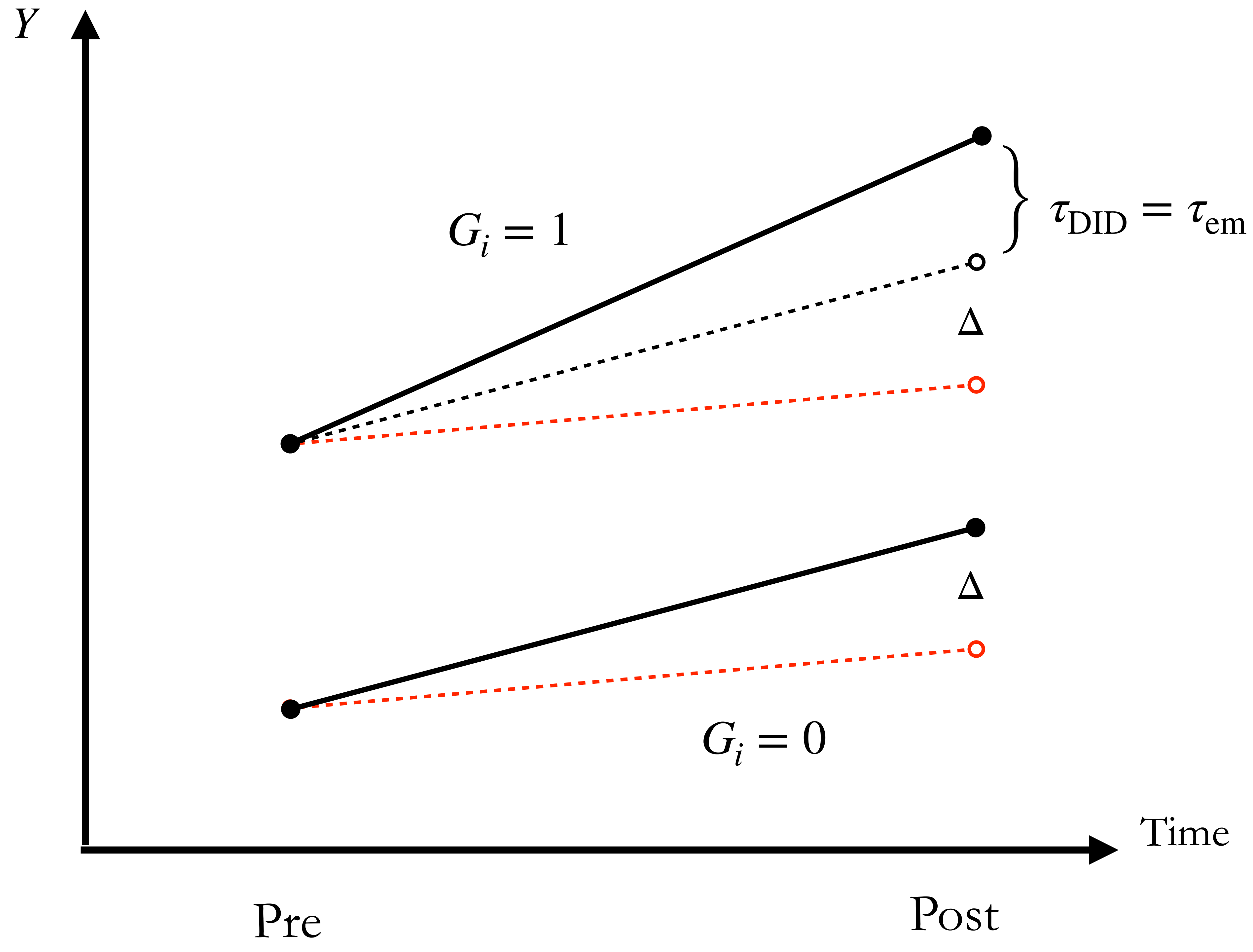
$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

$$\overset{?}{=}$$

$$\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0)] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0)]$$

Sufficient condition: $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$

Identifying Causal Interaction (Causal Moderation)



$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

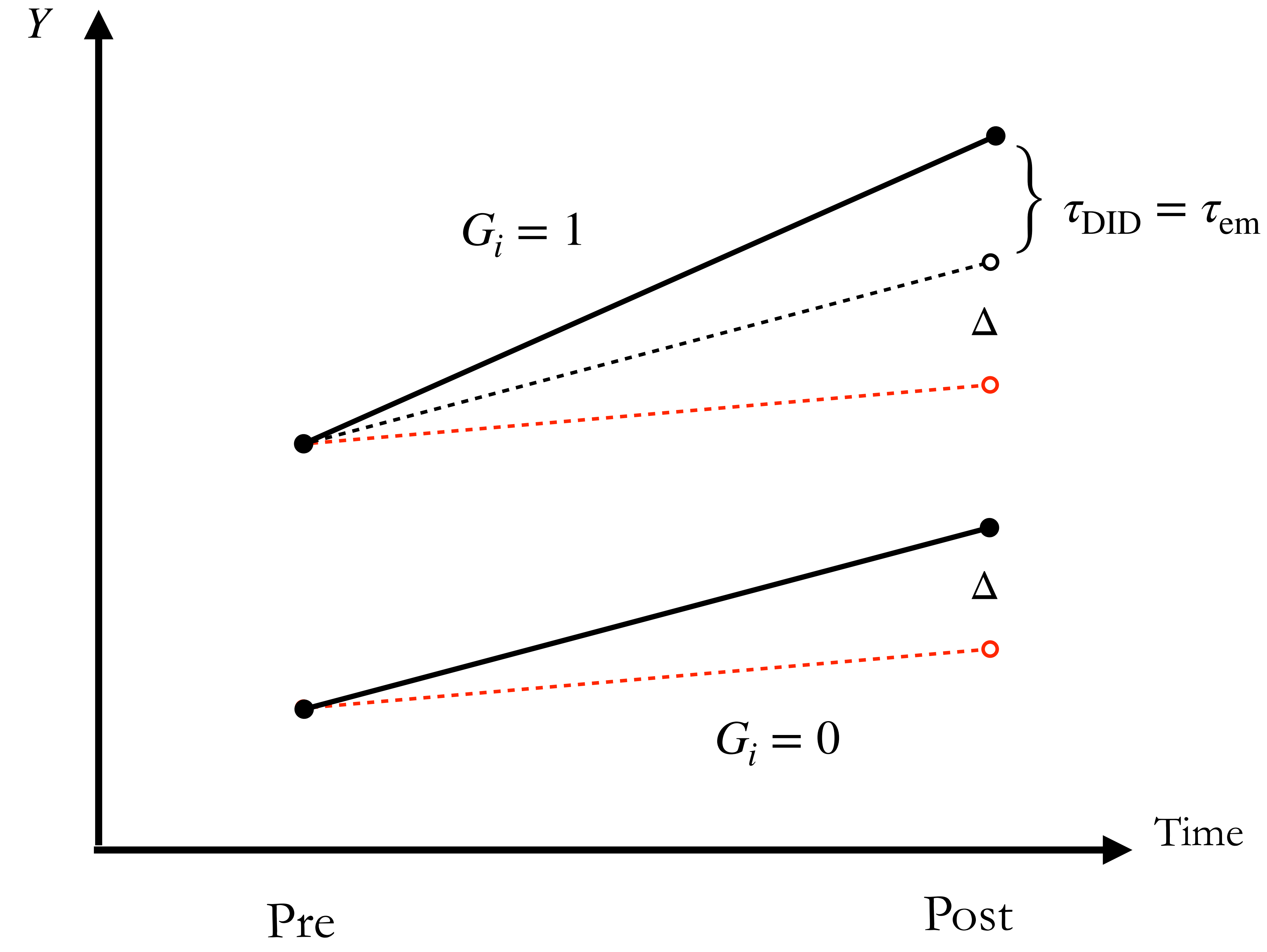
$$\stackrel{?}{=}$$

$$\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0)] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0)]$$

Sufficient condition: $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$

(Orthogonality between G and first-differenced potential outcomes)

Identifying Causal Interaction (Causal Moderation)



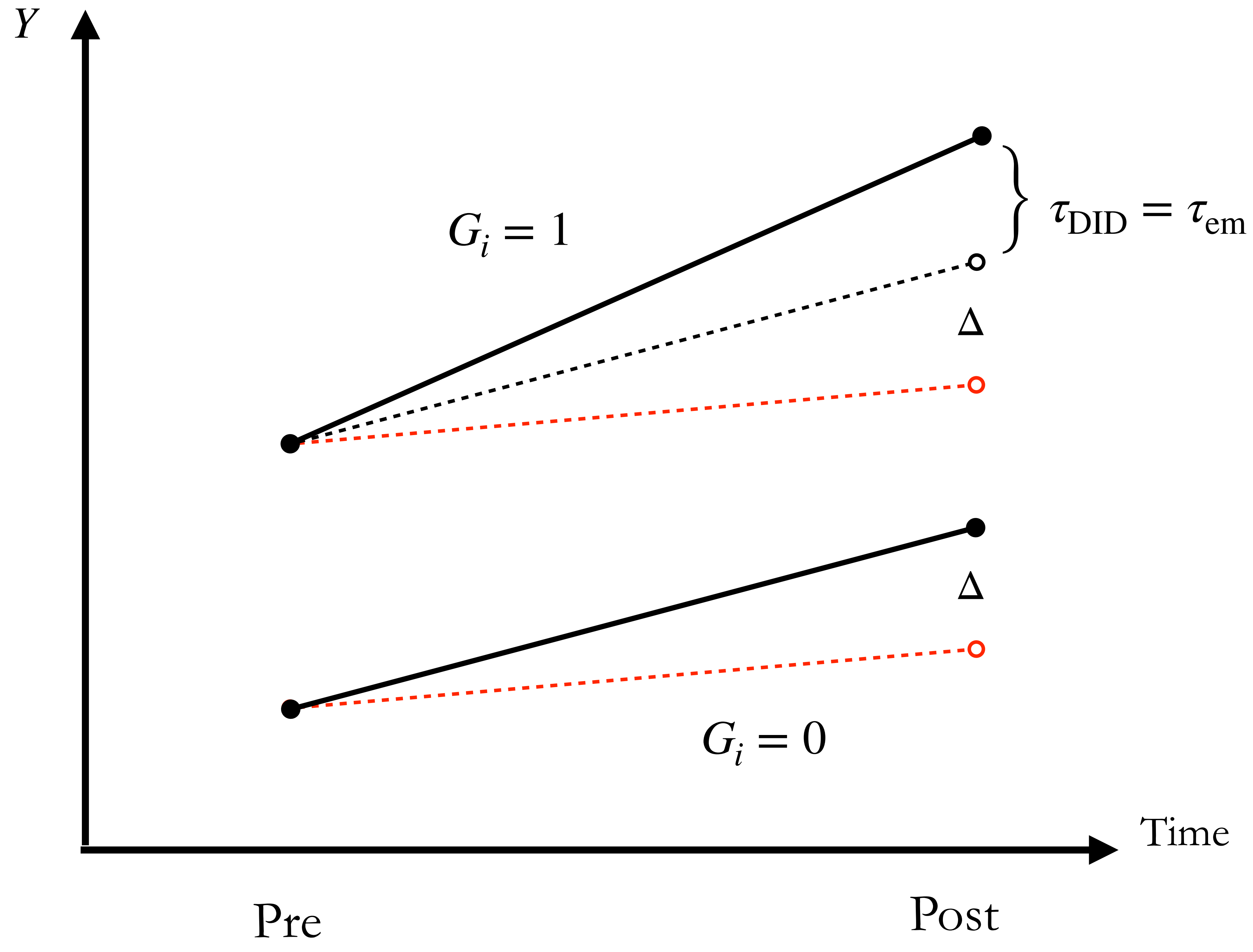
$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0] \overset{?}{=} \tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0)] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0)]$$

Sufficient condition: $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$

(Orthogonality between G and first-differenced potential outcomes)

Weaker than (quasi-)random assignment: $Y_{i,t}(g, z) \perp\!\!\!\perp G_i$

Identifying Causal Interaction (Causal Moderation)



$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0]$$

$$\overset{?}{=}$$

$$\tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0)] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0)]$$

Sufficient condition: $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$

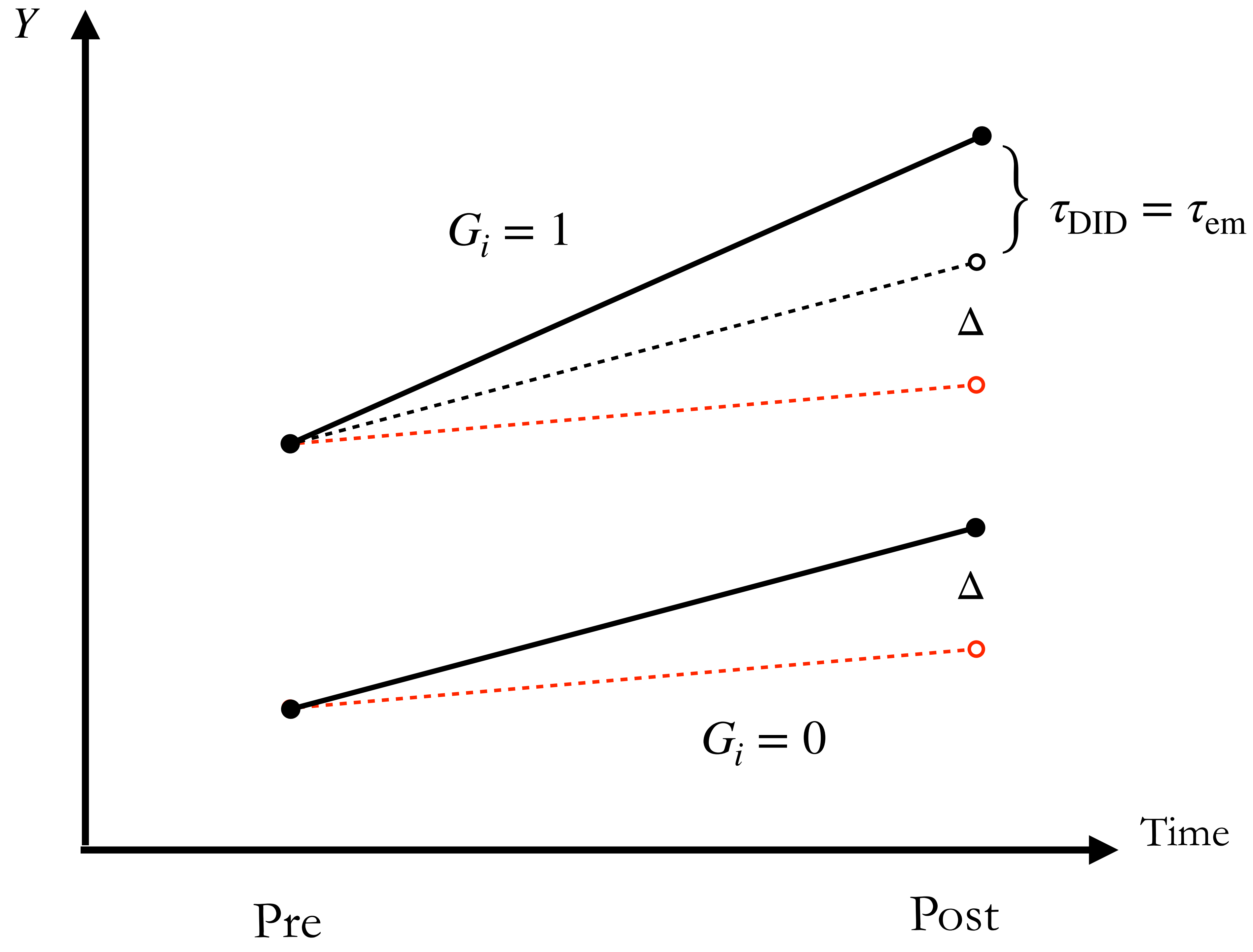
(Orthogonality between G and first-differenced potential outcomes)

Weaker than (quasi-)random assignment: $Y_{i,t}(g, z) \perp\!\!\!\perp G_i$

Exclusion Restriction on G

$Y_{i,\text{post}}(1,0) = Y_{i,\text{post}}(0,0)$ for all units with $Z_i = 0$

Identifying Causal Interaction (Causal Moderation)



$$\tau_{\text{DID}} = \tau_{\text{em}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0) \mid G_i = 1] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0) \mid G_i = 0] \overset{?}{=} \tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(1,0)] - \mathbb{E}[Y_{i,\text{post}}(0,1) - Y_{i,\text{post}}(0,0)]$$

Sufficient condition: $\Delta Y_i(g,z) \perp\!\!\!\perp G_i$

(Orthogonality between G and first-differenced potential outcomes)

Weaker than (quasi-)random assignment: $Y_{i,t}(g,z) \perp\!\!\!\perp G_i$

Exclusion Restriction on G

$Y_{i,\text{post}}(1,0) = Y_{i,\text{post}}(0,0)$ for all units with $Z_i = 0$

$$\Rightarrow \quad \tau_{\text{inter}} = \mathbb{E}[Y_{i,\text{post}}(1,1) - Y_{i,\text{post}}(0,1)] = \tau_{G|Z=1}$$

Summary of identification Results

τ_{DID}

A statistical estimand consistently estimated by $\hat{\tau}_{\text{DID}}$

No anticipation & Parallel trends

$=$

τ_{em}

An associative estimand describing effect heterogeneity

Exclusion restriction on Z for G = 0

$\|$

τ_{att}

A causal estimand targeting the effect of Z for G = 1

τ_{inter}

A causal estimand targeting causal moderation

$\tau_{G|Z=1}$

A causal estimand targeting the effect of G when Z = 1

Summary of identification Results

τ_{DID} — No anticipation & Parallel trends — τ_{em} — $\Delta Y(g,z) \perp\!\!\!\perp G$ — τ_{inter}

τ_{DID}: A statistical estimand consistently estimated by $\hat{\tau}_{\text{DID}}$

τ_{em}: An associative estimand describing effect heterogeneity

τ_{inter}: A causal estimand targeting causal moderation

Exclusion restriction on Z for G = 0

τ_{att}: A causal estimand targeting the effect of Z for G = 1

$\tau_{G|Z=1}$: A causal estimand targeting the effect of G when Z = 1

Summary of identification Results

$$\tau_{\text{DID}} \overset{\text{No anticipation \& Parallel trends}}{=\!=} \tau_{\text{em}} \overset{\Delta Y(g,z) \perp\!\!\!\perp G}{=\!=} \tau_{\text{inter}}$$

τ_{DID}: A statistical estimand consistently estimated by $\hat{\tau}_{\text{DID}}$

τ_{em}: An associative estimand describing effect heterogeneity

τ_{inter}: A causal estimand targeting causal moderation

Exclusion restriction on Z for G = 0: $\tau_{\text{em}} = \tau_{\text{att}}$

Exclusion restriction on G when Z = 0: $\tau_{\text{inter}} = \tau_{G|Z=1}$

τ_{att}: A causal estimand targeting the effect of Z for G = 1

$\tau_{G|Z=1}$: A causal estimand targeting the effect of G when Z = 1

Canonical vs Factorial DID, Two Research Designs

Canonical vs Factorial DID, Two Research Designs

Canonical DID (original)

Define canonical DID research design as the combination of:

- 2×2 data structure
- Identification results:
  - Under no anticipation & parallel trends, $\hat{\tau}_{\text{DID}}$ identifies τ_{att}

Canonical vs Factorial DID, Two Research Designs

Canonical DID (reframed)

Define canonical DID research design as the combination of:

- 2×2 data structure & universal exposure
- Identification results:
  - Under no anticipation & parallel trends & exclusion restriction on Z, $\hat{\tau}_{\text{DID}}$ identifies τ_{att}

Canonical vs Factorial DID, Two Research Designs

Canonical DID (reframed)

Define canonical DID research design as the combination of:

- 2×2 data structure & universal exposure
- Identification results:
  - Under no anticipation & parallel trends & exclusion restriction on Z, $\hat{\tau}_{\text{DID}}$ identifies τ_{att}

Factorial DID

Define factorial DID research design as the combination of:

- 2×2 data structure & universal exposure
- Identification results:
  1. Under no anticipation & parallel trends, $\hat{\tau}_{\text{DID}}$ identifies τ_{em}
  2. Under no anticipation & parallel trends & $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$, $\hat{\tau}_{\text{DID}}$ identifies τ_{inter}
  3. Under no anticipation & parallel trends & $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$ & exclusion restriction on G, $\hat{\tau}_{\text{DID}}$ identifies $\tau_{G|Z=1}$

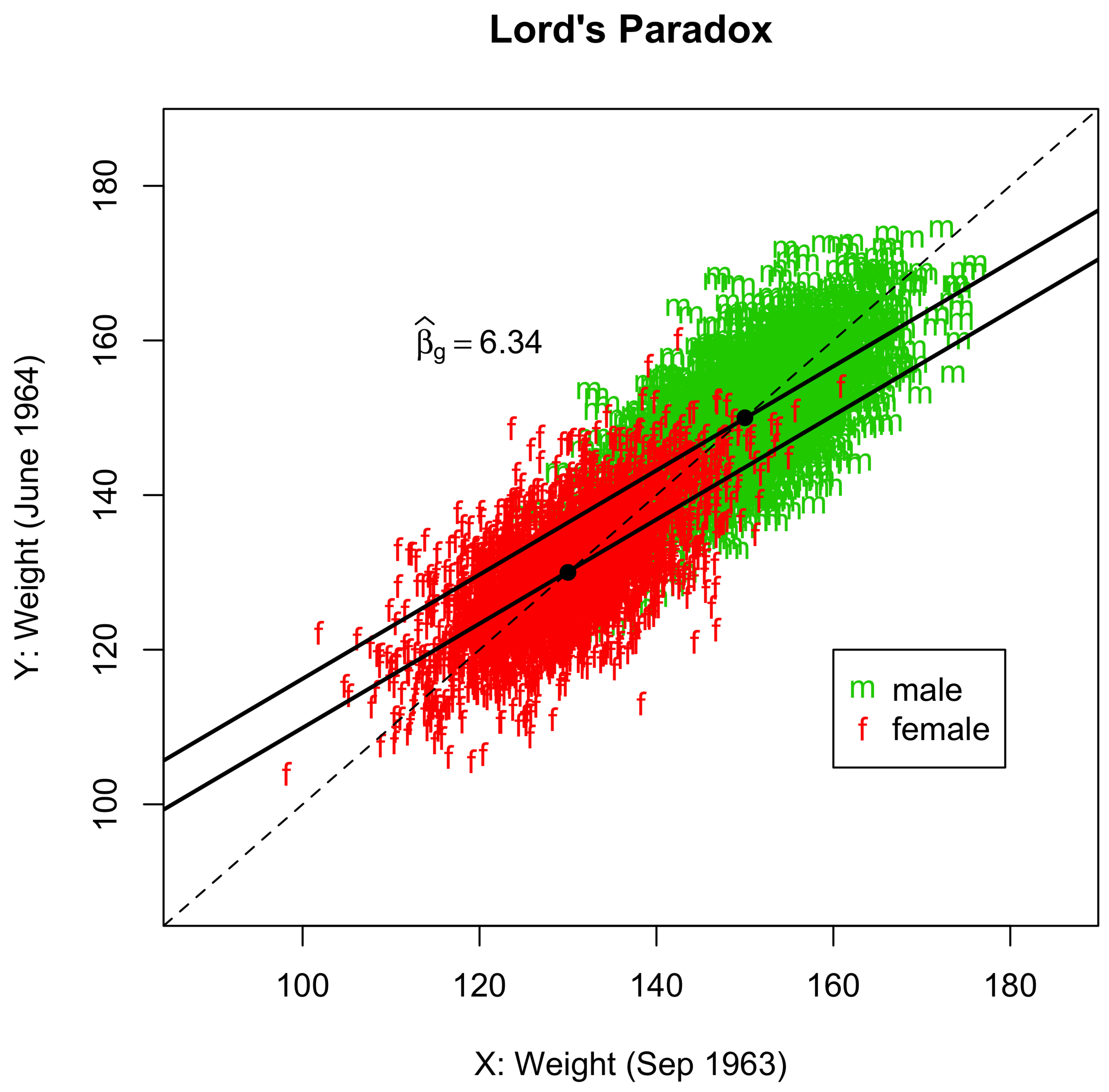
Lord's Paradox

Lord's Paradox

"A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects ... [t]he weight of each student at the time of his (/her) arrival in September and his weight the following June are recorded." (Lord 1967, p. 304)

Lord’s Paradox

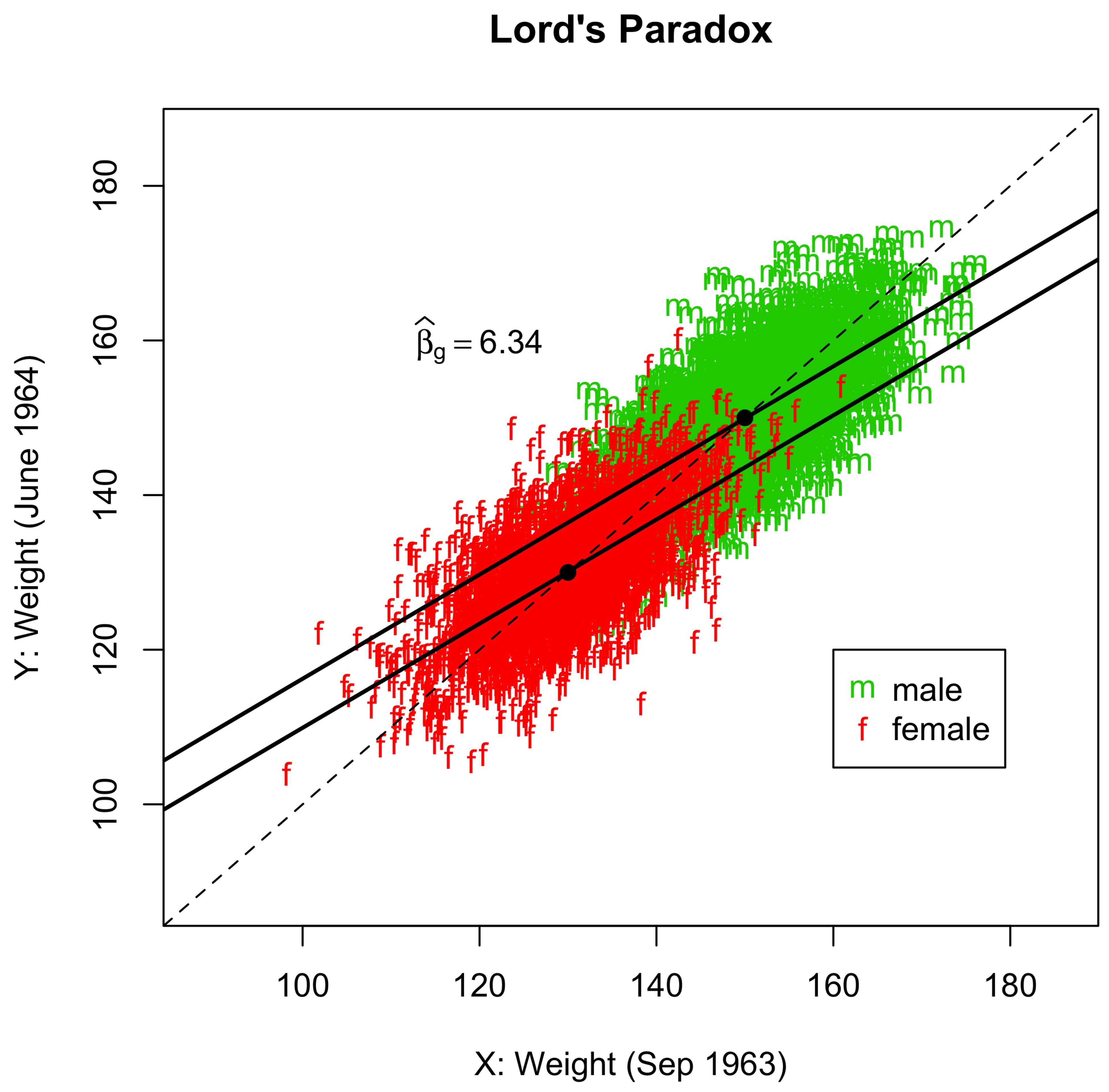
“A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects … [t]he weight of each student at the time of his (/her) arrival in September and his weight the following June are recorded.” (Lord 1967, p. 304)



Lord’s Paradox

“A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects ... [t]he weight of each student at the time of his (/her) arrival in September and his weight the following June are recorded.” (Lord 1967, p. 304)

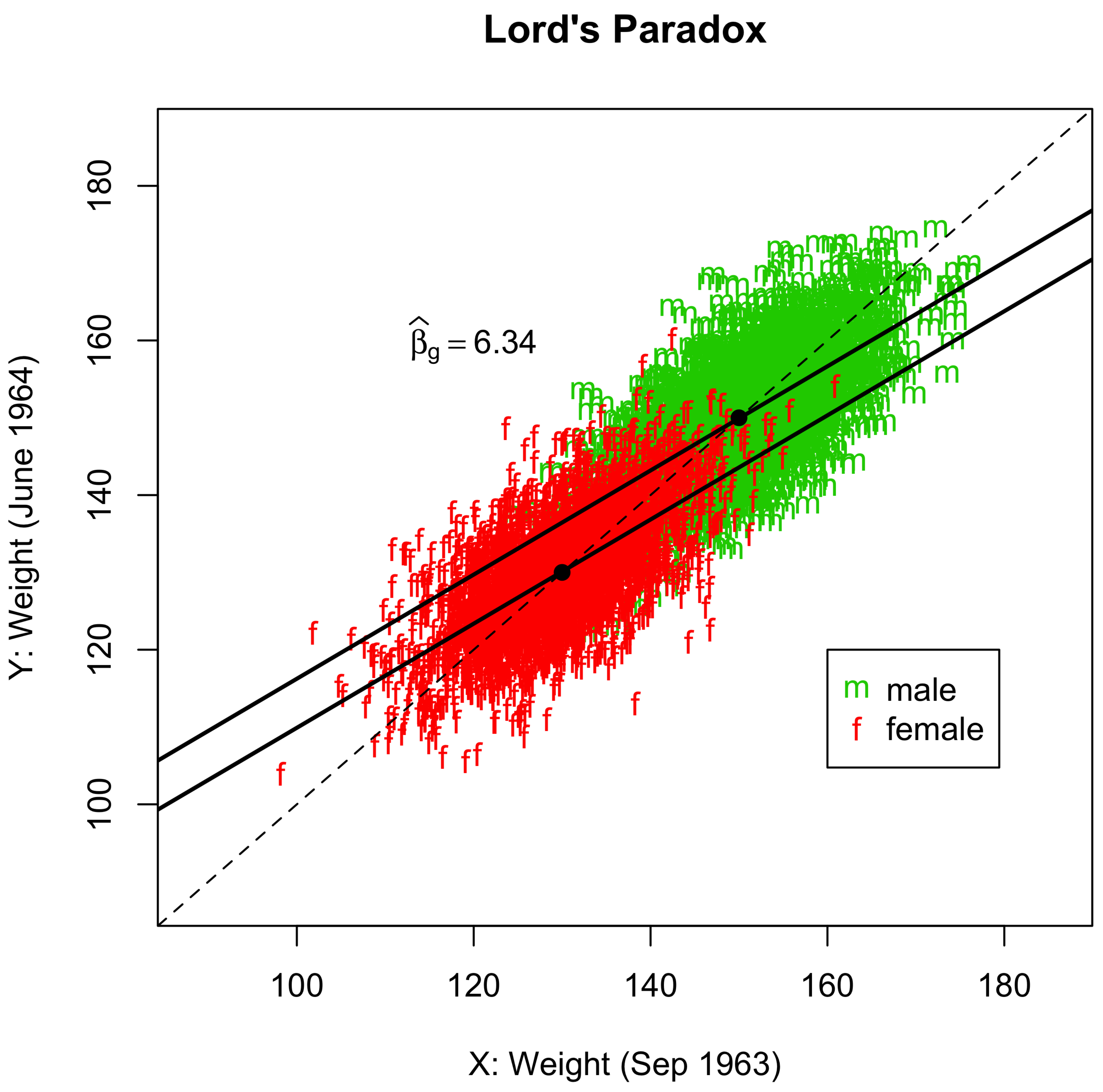
- Statistician 1: zero effect for both groups because average weight does not changeeitherbgroup



Lord's Paradox

"A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects ... [t]he weight of each student at the time of his (/her) arrival in September and his weight the following June are recorded." (Lord 1967, p. 304)

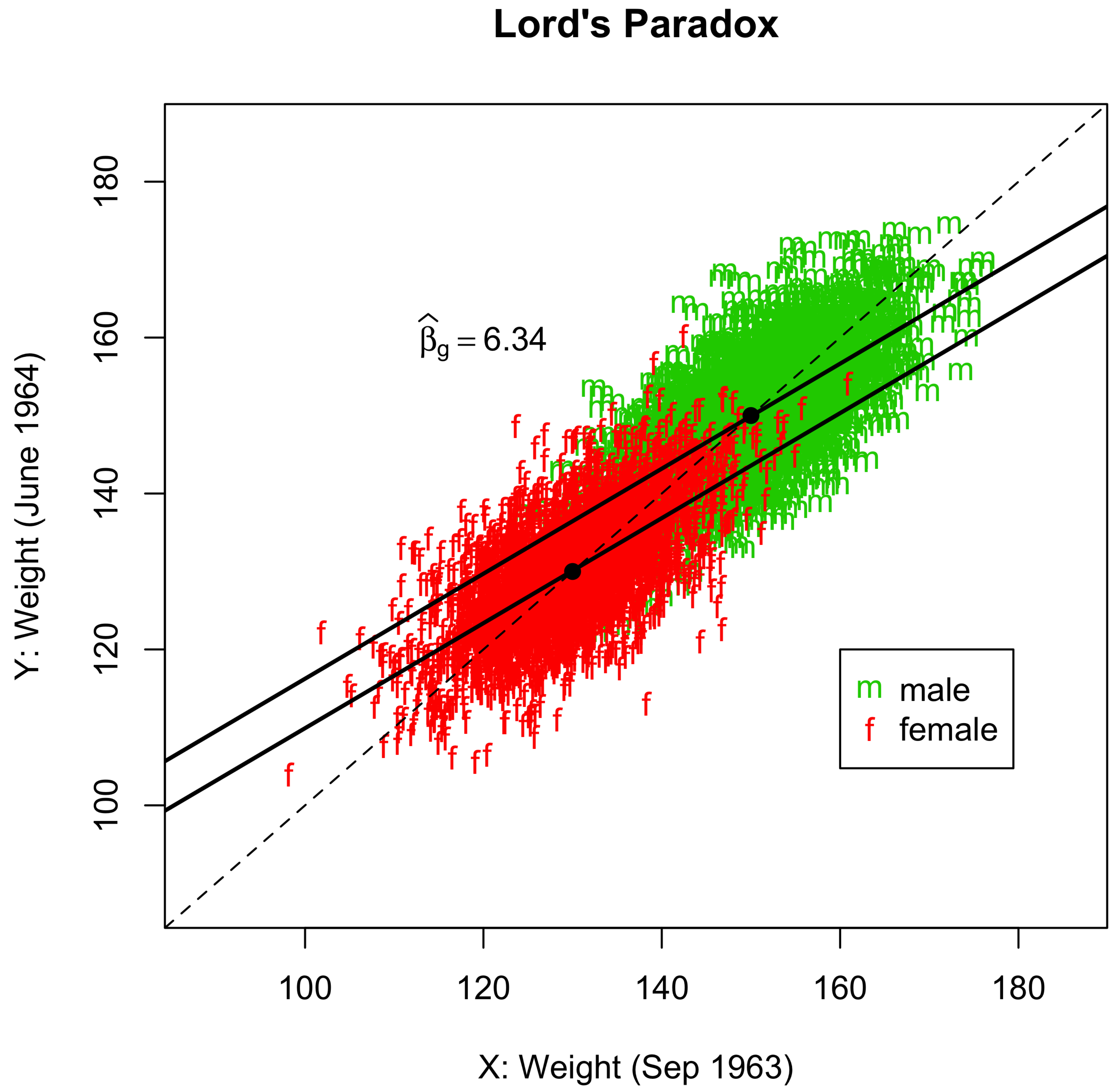
- Statistician 1: zero effect for both groups because average weight does not changeeitherbgroup
- Statistician 2: larger effect on male students (6.34) from: $Y_{i,\text{post}} \sim 1 + \ male_i + Y_{i,\text{pre}}$



Lord's Paradox

"A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex differences in these effects ... [t]he weight of each student at the time of his (/her) arrival in September and his weight the following June are recorded." (Lord 1967, p. 304)

- Statistician 1: zero effect for both groups because average weight does not changeeitherbgroup
- Statistician 2: larger effect on male students (6.34) from: $Y_{i,\text{post}} \sim 1 + \ male_i + Y_{i,\text{pre}}$
- Resolution
  - Holland and Rubin (1986): Untestable assumptions on $Y_{i,\text{post}}(g,0)$
  - Statistician 1: $Y_{i,\text{post}}(g,0) = Y_{i,\text{pre}}$
  - Statistician 2: $Y_{i,\text{post}}(g,0) = \beta Y_{i,\text{pre}} + \gamma_g$
  - Statistician 3 (Factorial DID): Under no anticipation & parallel trends, $\tau_{em} = \tau_{\text{DID}} = 0$



Roadmap

- Motivation
- Setup & Estimands
- Identification
- Extensions
  - Conditionally valid assumptions
  - *Multiple pre- and post- periods
  - *Multi-valued G
- Example: Clans and Calamity

Extension to Conditionally Valid Assumptions

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

  - Believe $\Delta Y_i(g, z) \perp\!\!\!\perp G_i \mid X_i$ is more plausible than $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

  - Believe $\Delta Y_i(g, z) \perp\!\!\!\perp G_i \mid X_i$ is more plausible than $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$
  - Simple improvements: (a) demean X_i; (b) add interaction terms

$$Y_{it} = \alpha_i + \xi_t \ + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \gamma^{\mathrm{T}} \mathrm{G_i} \cdot \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

  - Believe $\Delta Y_i(g, z) \perp\!\!\!\perp G_i \mid X_i$ is more plausible than $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$
  - Simple improvements: (a) demean X_i; (b) add interaction terms

$$Y_{it} = \alpha_i + \xi_t \ + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \gamma^{\mathrm{T}} \mathrm{G_i} \cdot \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

- Can leverage more flexible models of ΔY_i on X_i for subgroups $G_i = 1{,}0$

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

  - Believe $\Delta Y_i(g, z) \perp\!\!\!\perp G_i \mid X_i$ is more plausible than $\Delta Y_i(g, z) \perp\!\!\!\perp G_i$
  - Simple improvements: (a) demean X_i; (b) add interaction terms

$$Y_{it} = \alpha_i + \xi_t \ + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \gamma^{\mathrm{T}} \mathrm{G_i} \cdot \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

- Can leverage more flexible models of ΔY_i on X_i for subgroups $G_i = 1{,}0$
  - Transform data into wide form; replace Y_i with ΔY_i

Extension to Conditionally Valid Assumptions

- Researchers often run the following TWFE regression:

$$Y_{it} = \alpha_i + \xi_t + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

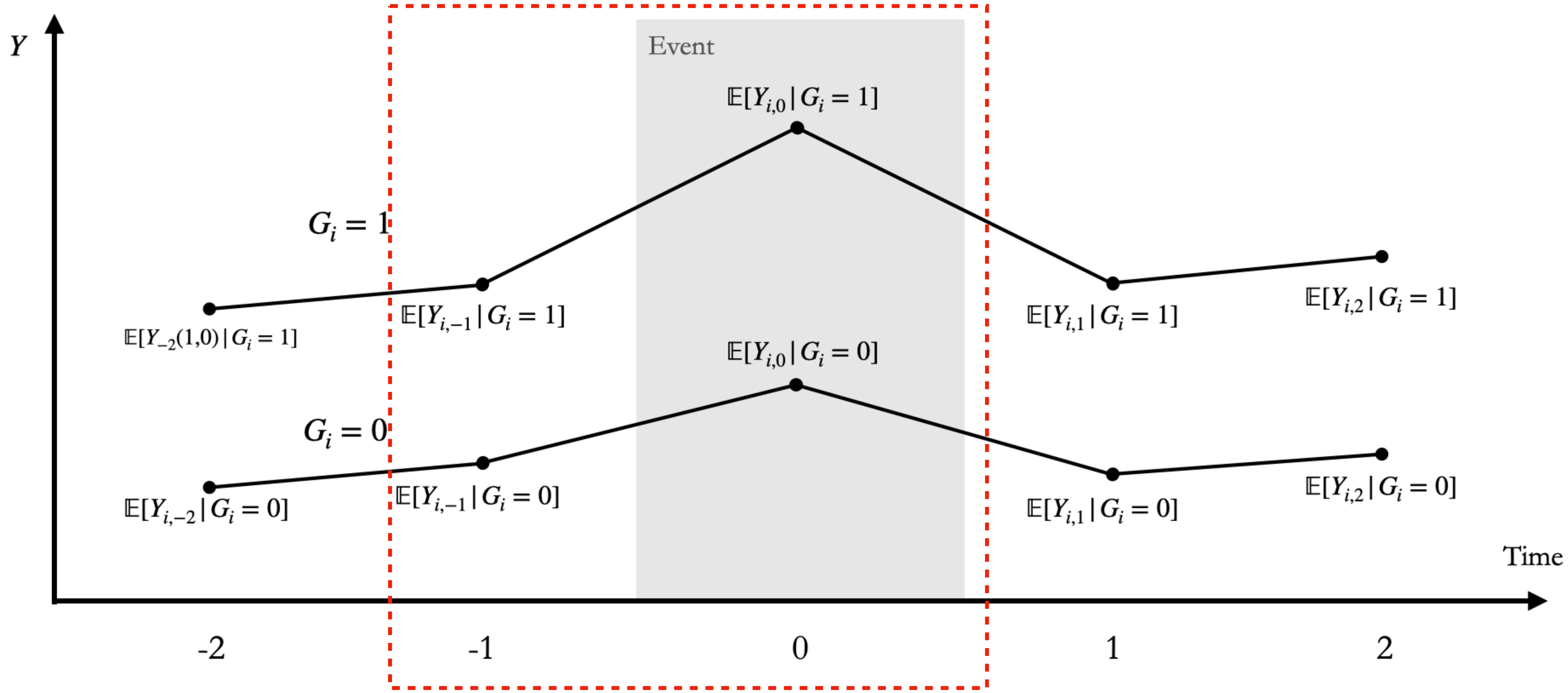
  - Believe $\Delta Y_i(g,z) \perp\!\!\!\perp G_i \mid X_i$ is more plausible than $\Delta Y_i(g,z) \perp\!\!\!\perp G_i$
  - Simple improvements: (a) demean X_i; (b) add interaction terms

$$Y_{it} = \alpha_i + \xi_t \ + \tau G_i \cdot \mathrm{Post_t} + \beta^{\mathrm{T}} \mathrm{X_i} \cdot \mathrm{Post_t} + \gamma^{\mathrm{T}} \mathrm{G_i} \cdot \mathrm{X_i} \cdot \mathrm{Post_t} + \epsilon_{\mathrm{it}}$$

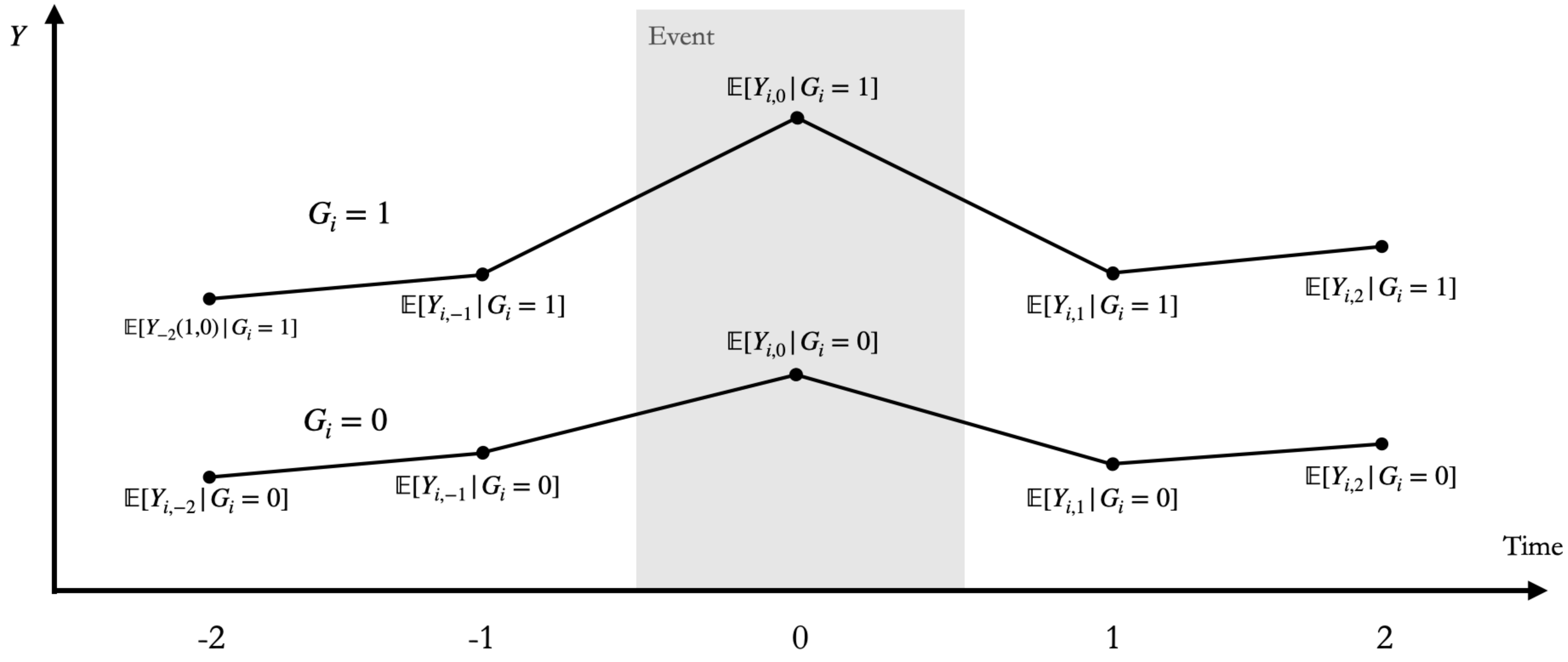
- Can leverage more flexible models of ΔY_i on X_i for subgroups $G_i = 1{,}0$
  - Transform data into wide form; replace Y_i with ΔY_i
  - Apply a variety of estimators developed for selection-on-observables designs

    (e.g., stratification, matching, balancing, IPW, AIPW, outcome modeling, double machine learning...)

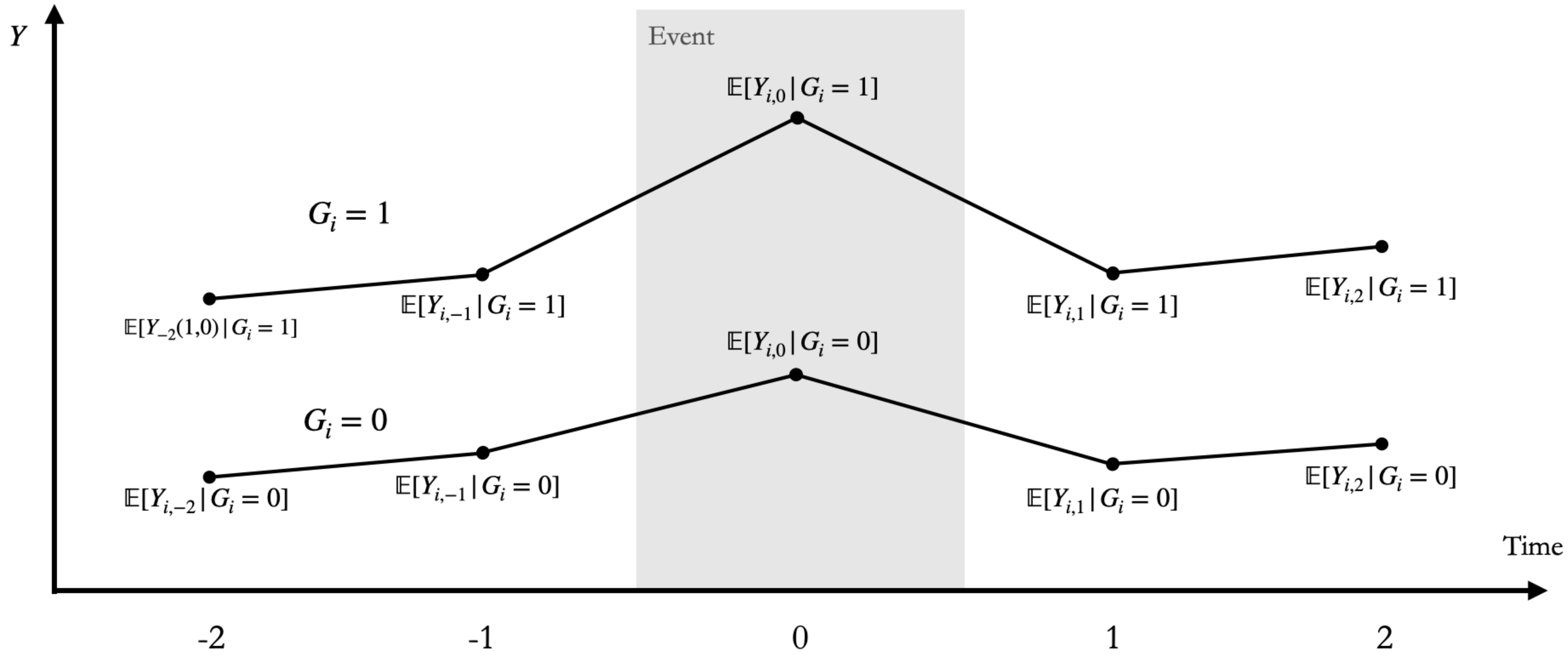
Extension to Multiple Pre- and Post-Periods



Extension to Multiple Pre- and Post-Periods

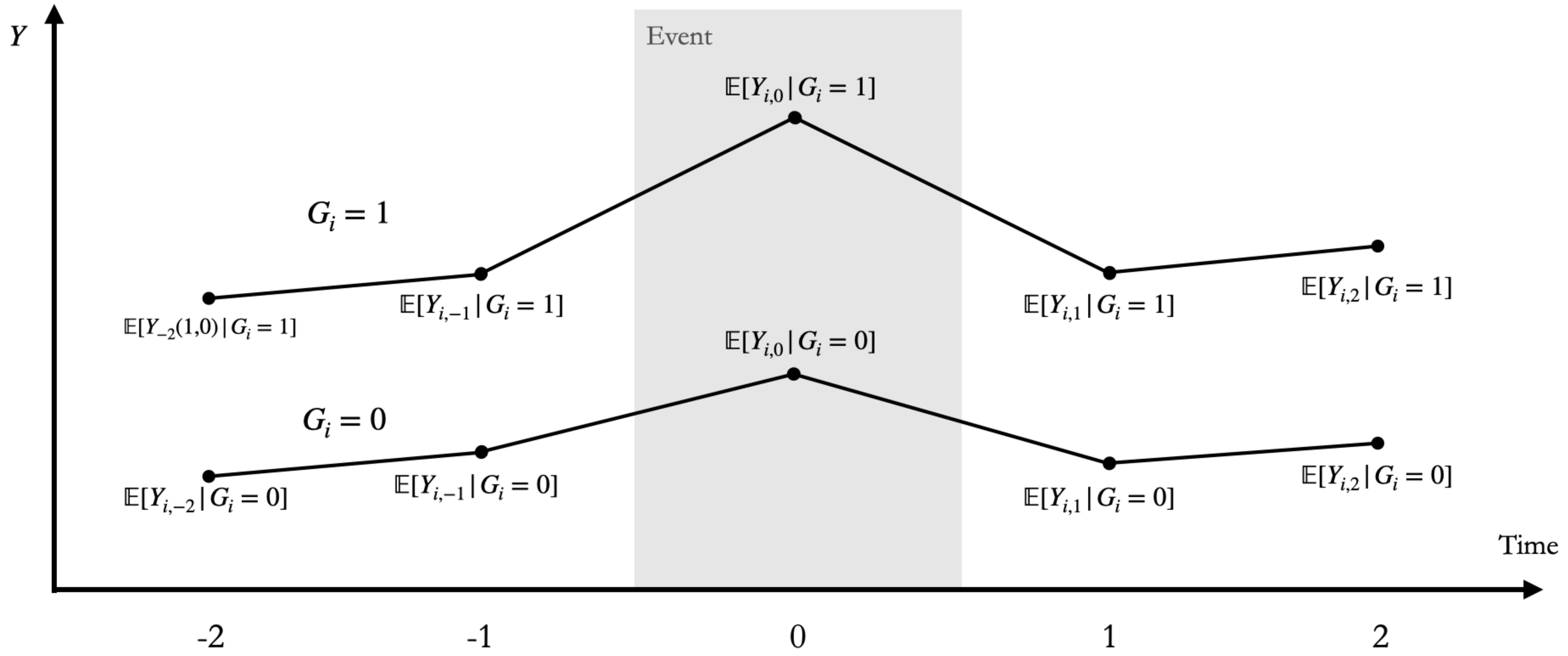


Extension to Multiple Pre- and Post-Periods



- Lesson 1: Pretrend tests can help assess the parallel trends assumption, but *not* $\Delta Y(g, z) \perp\!\!\!\perp G$

Extension to Multiple Pre- and Post-Periods

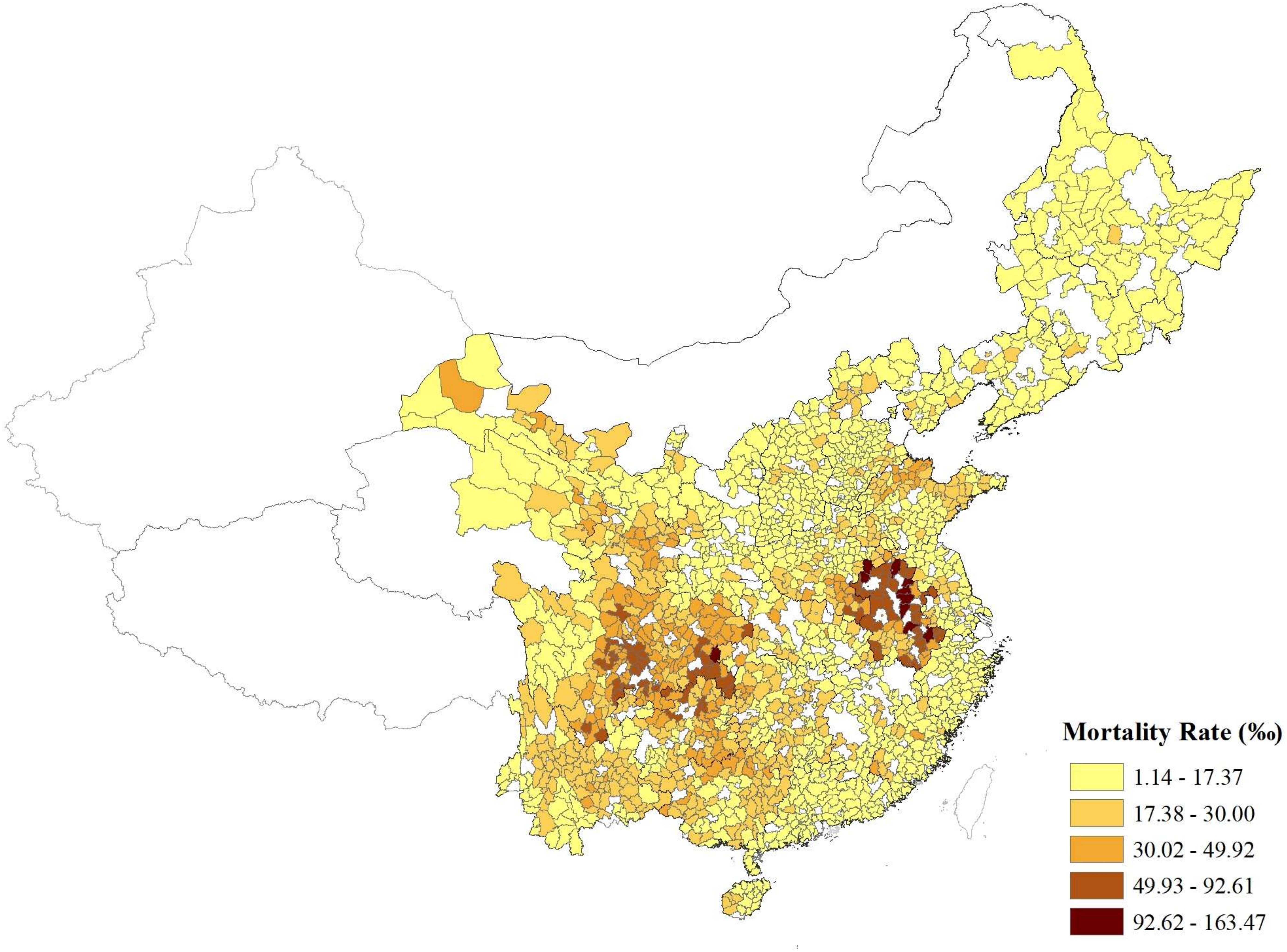


- Lesson 1: Pretrend tests can help assess the parallel trends assumption, but *not* $\Delta Y(g,z) \perp\!\!\!\perp G$
- Lesson 2: Using post-periods as non-event periods requires an additional “no carryover effect” assumption

Roadmap

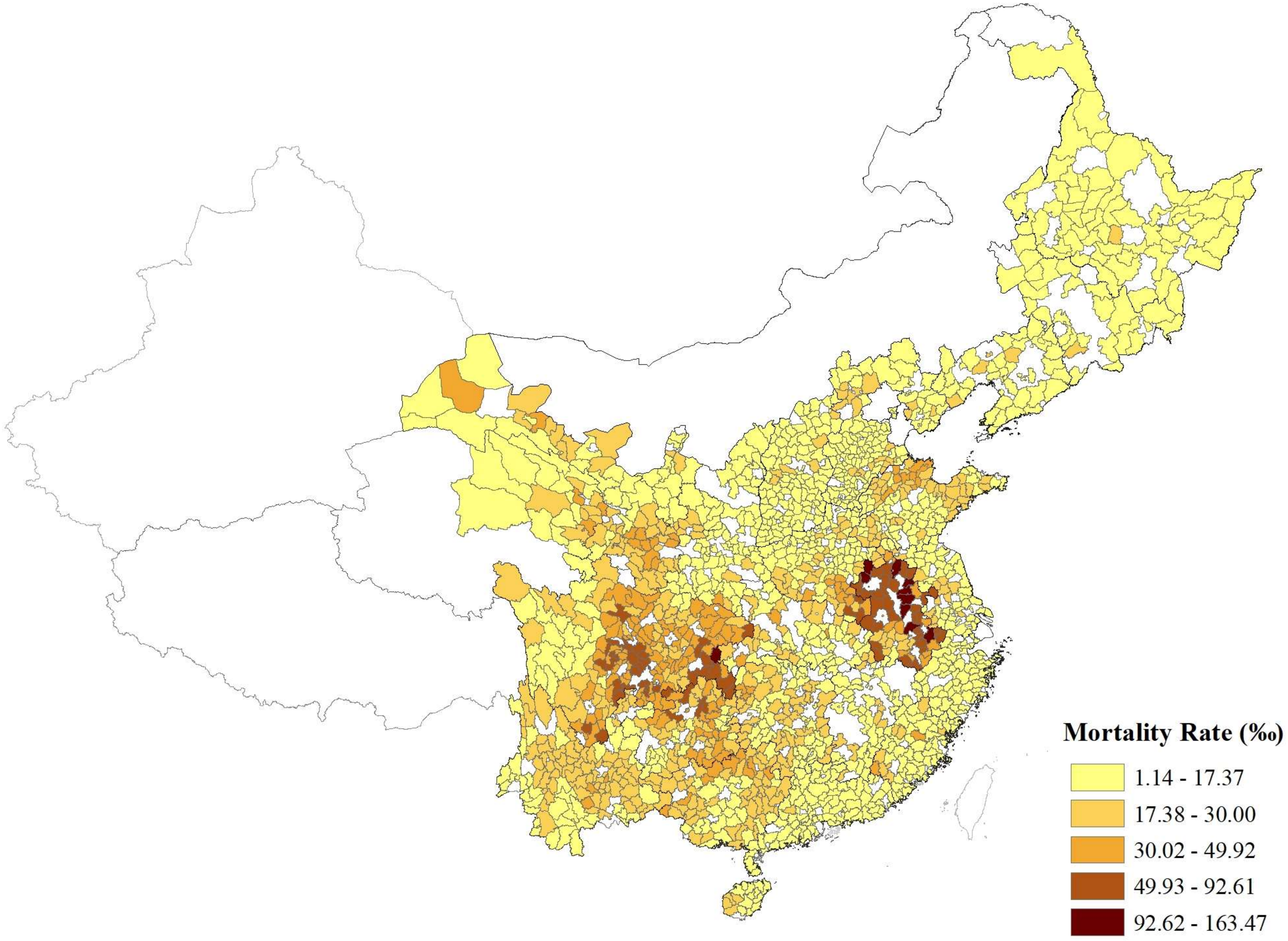
- Motivation
- Setup & Estimands
- Identification
- Extensions
- Example: Clans and Calamity

Example: Clans and Calamity



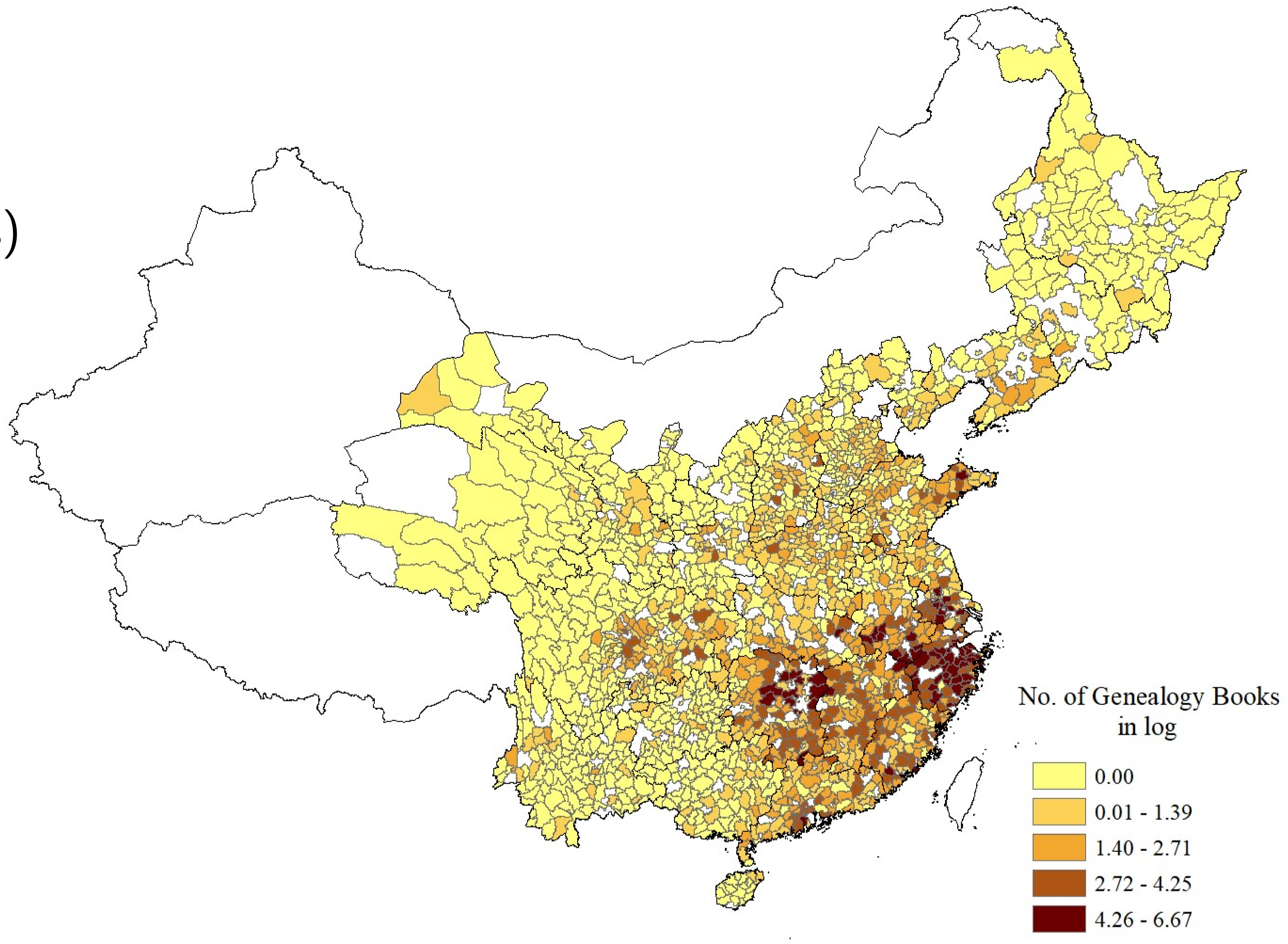
Example: Clans and Calamity

- Event — The Great famine (1958-1961)

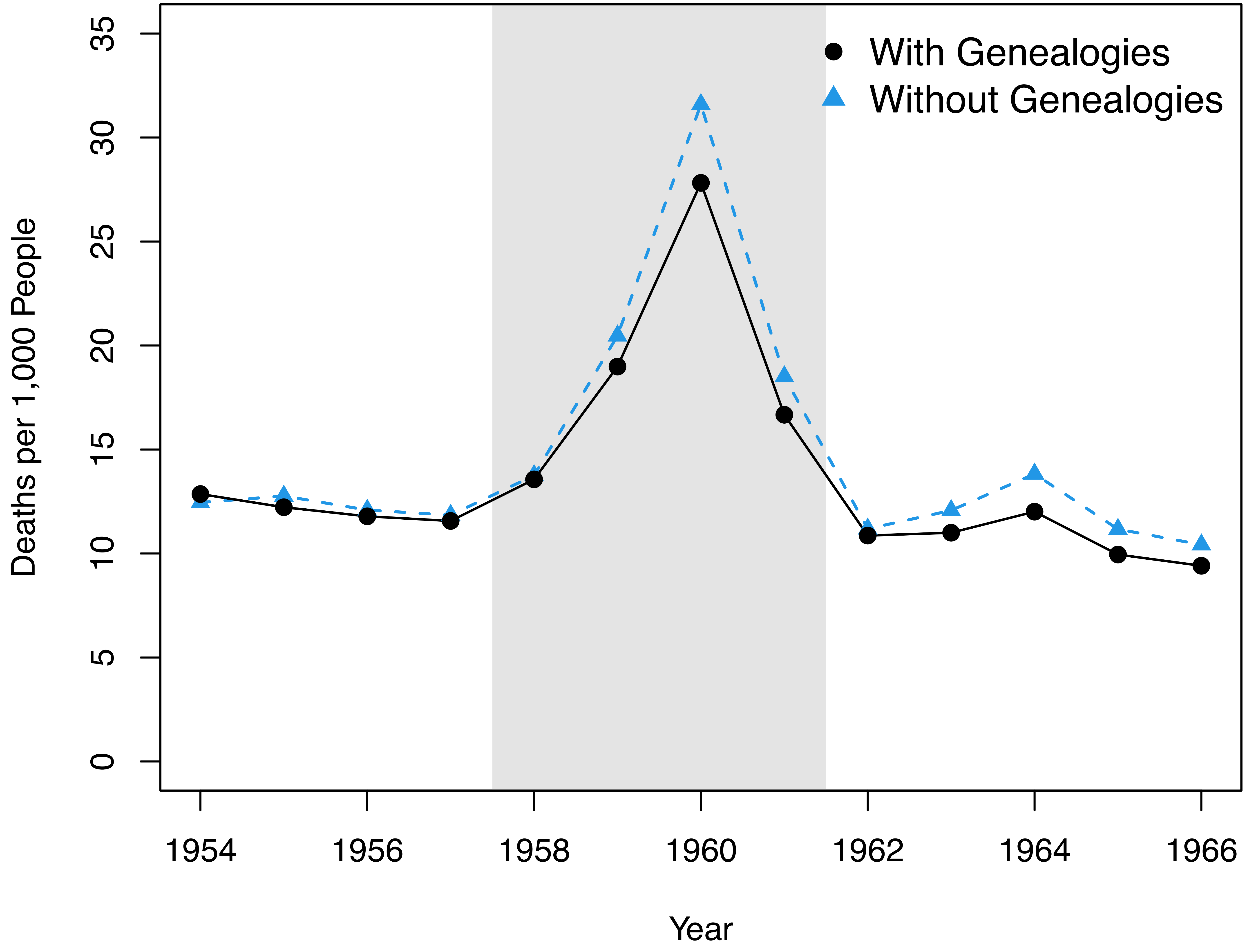


Example: Clans and Calamity

- Event — The Great famine (1958-1961)
- G — Social Capital (proxied by genealogies)
  - No Genealogies: 412 counties
  - Have pre-PRC genealogies: 509 counties



Raw Data: Group Means



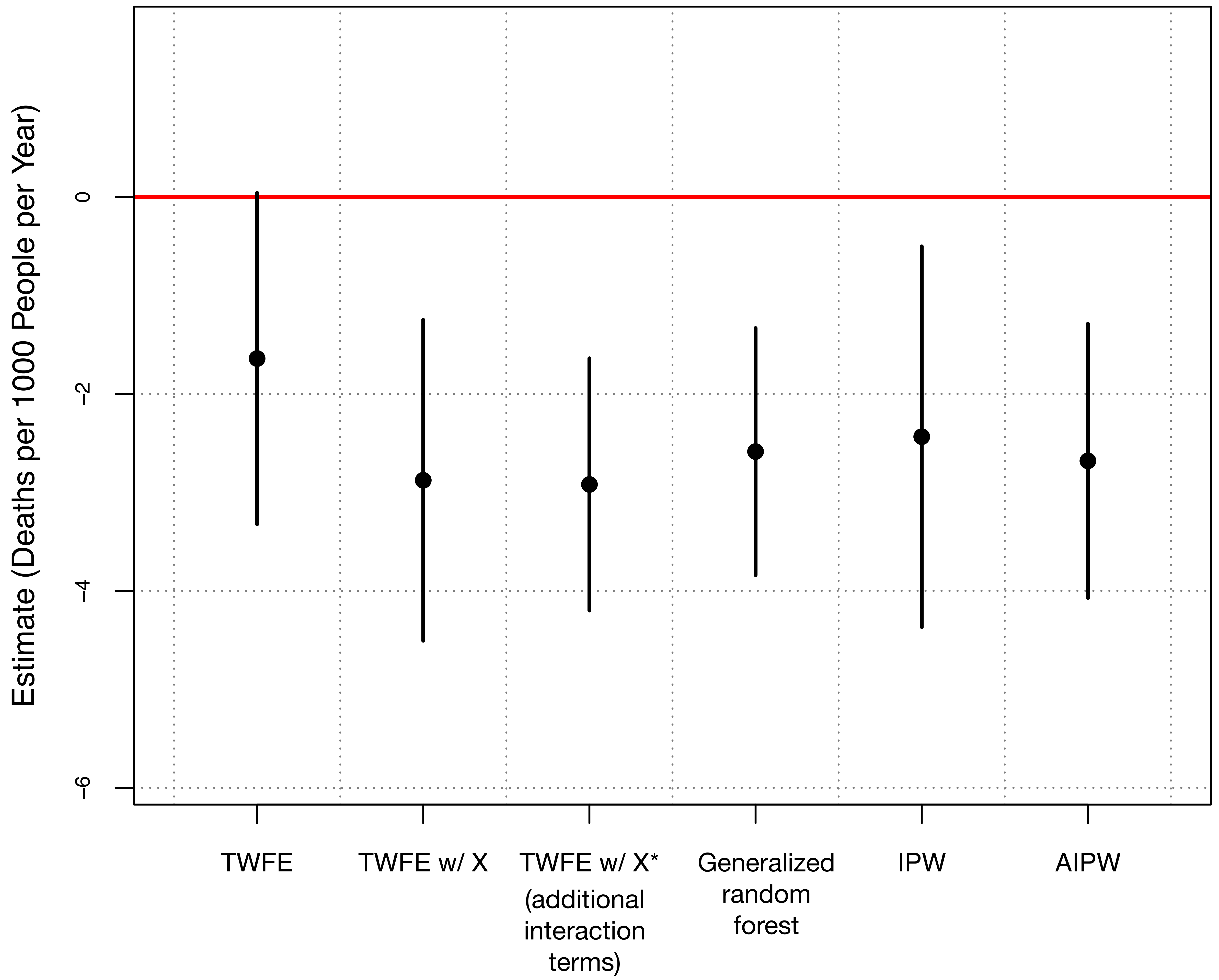
Conditional on Covariates

Conditional on Covariates

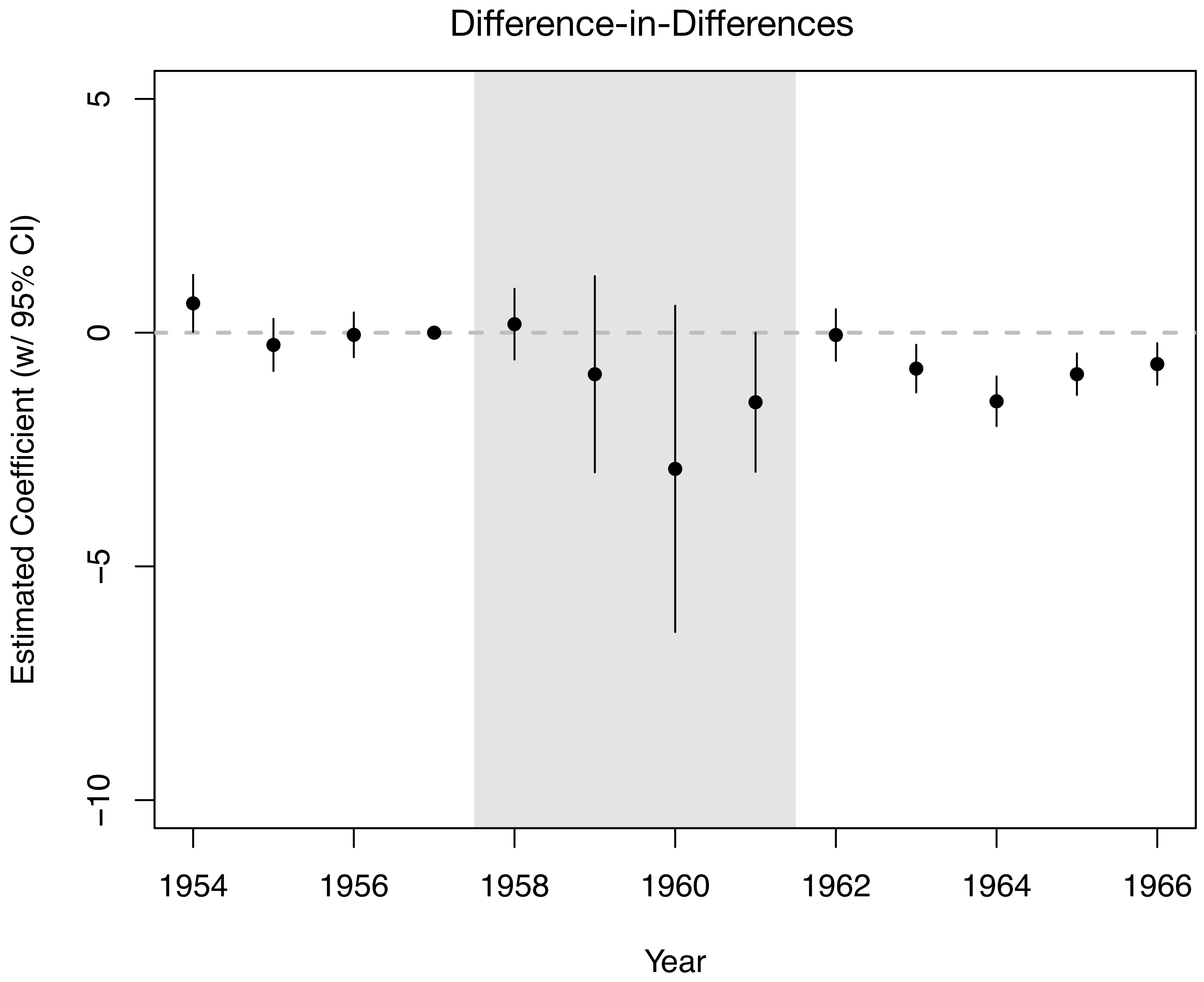
- Pre-event Covariates
  - Log population size
  - Per capita grain production
  - Ratio of non-farming land
  - Urbanization ratio
  - Distance from Beijing
  - Distance from the provincial capital
  - Share of ethnic minorities
  - Rice suitability
  - Average years of education

Conditional on Covariates

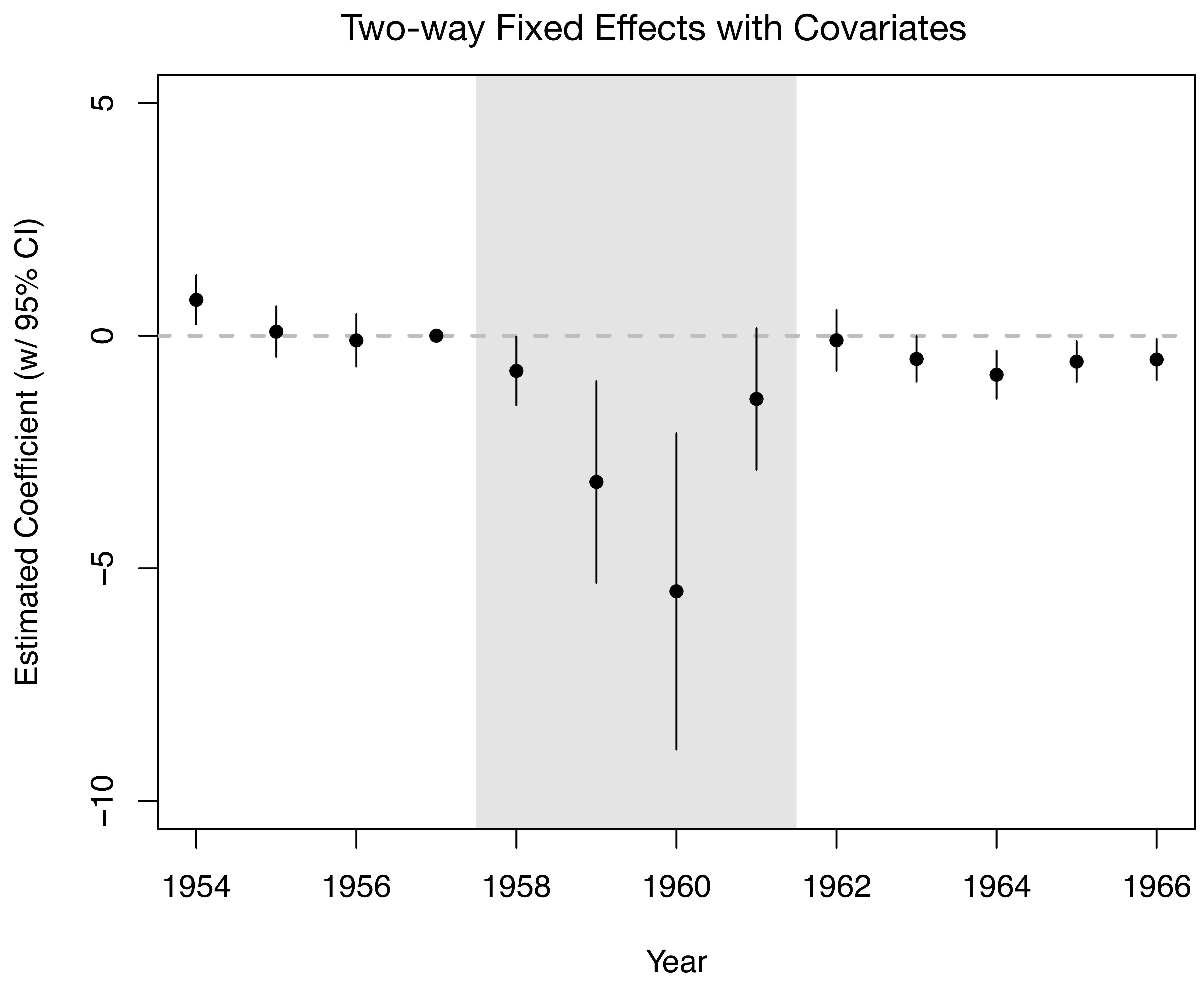
- Pre-event Covariates
  - Log population size
  - Per capita grain production
  - Ratio of non-farming land
  - Urbanization ratio
  - Distance from Beijing
  - Distance from the provincial capital
  - Share of ethnic minorities
  - Rice suitability
  - Average years of education



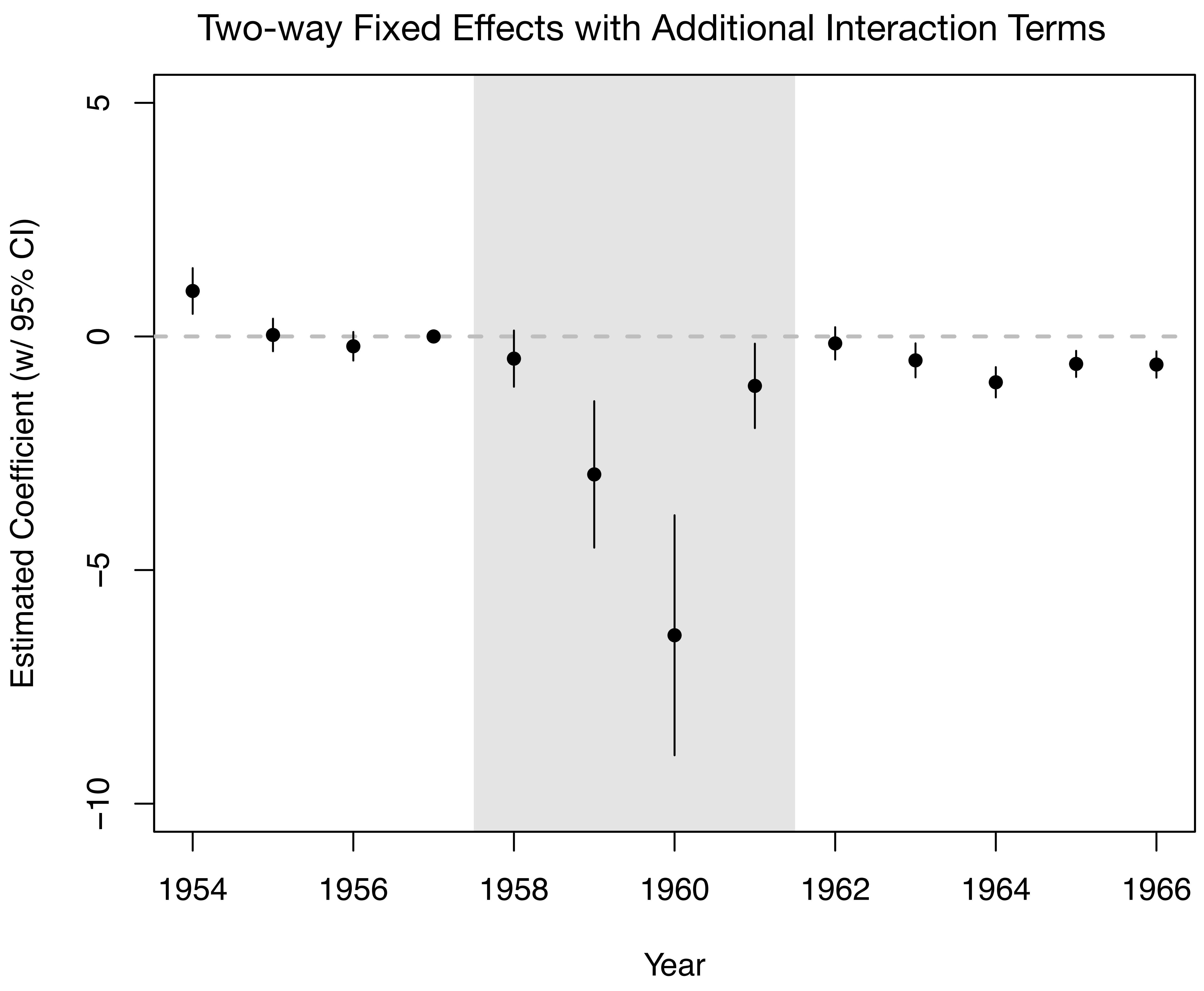
Dynamic Estimates



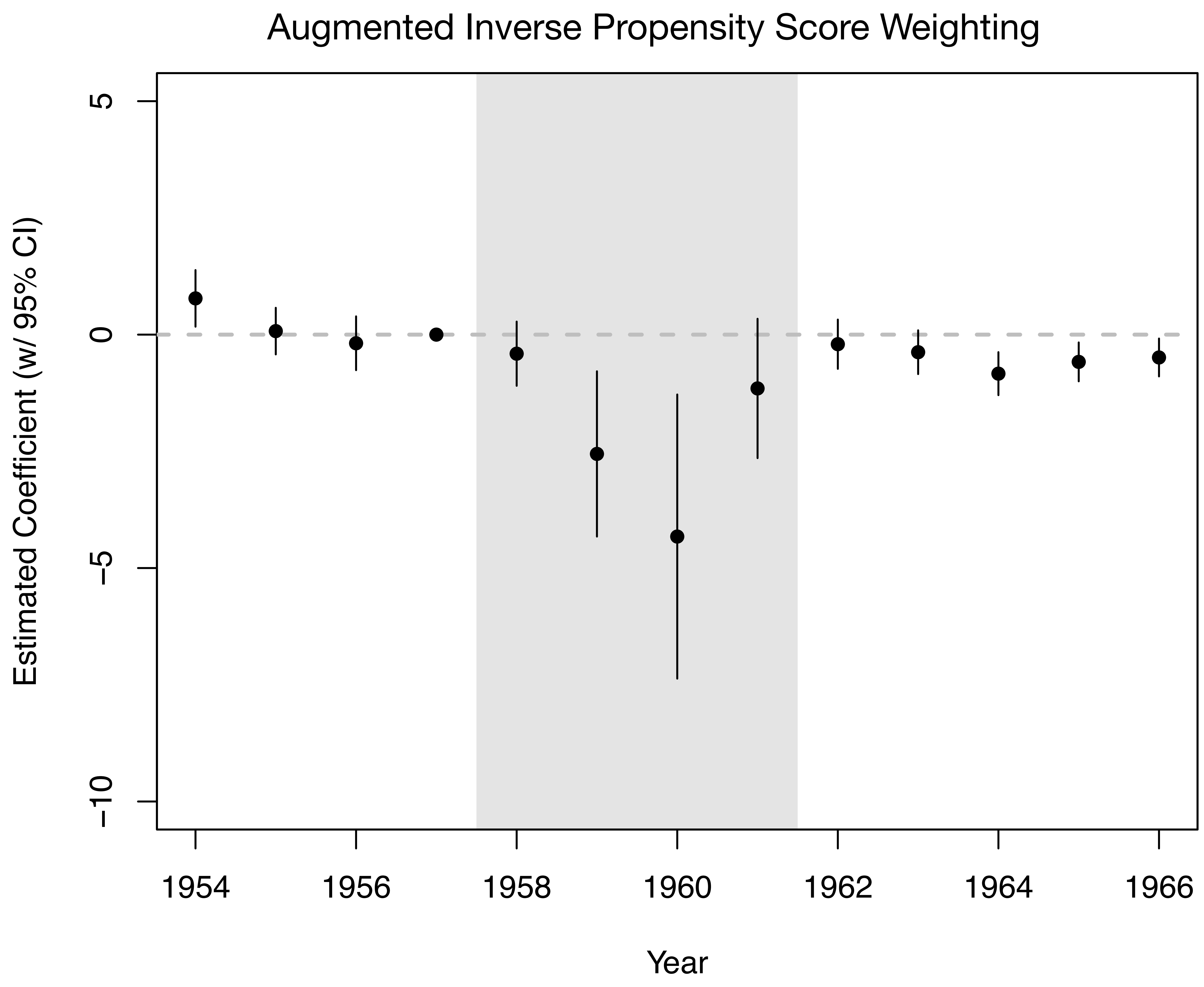
Dynamic Estimates



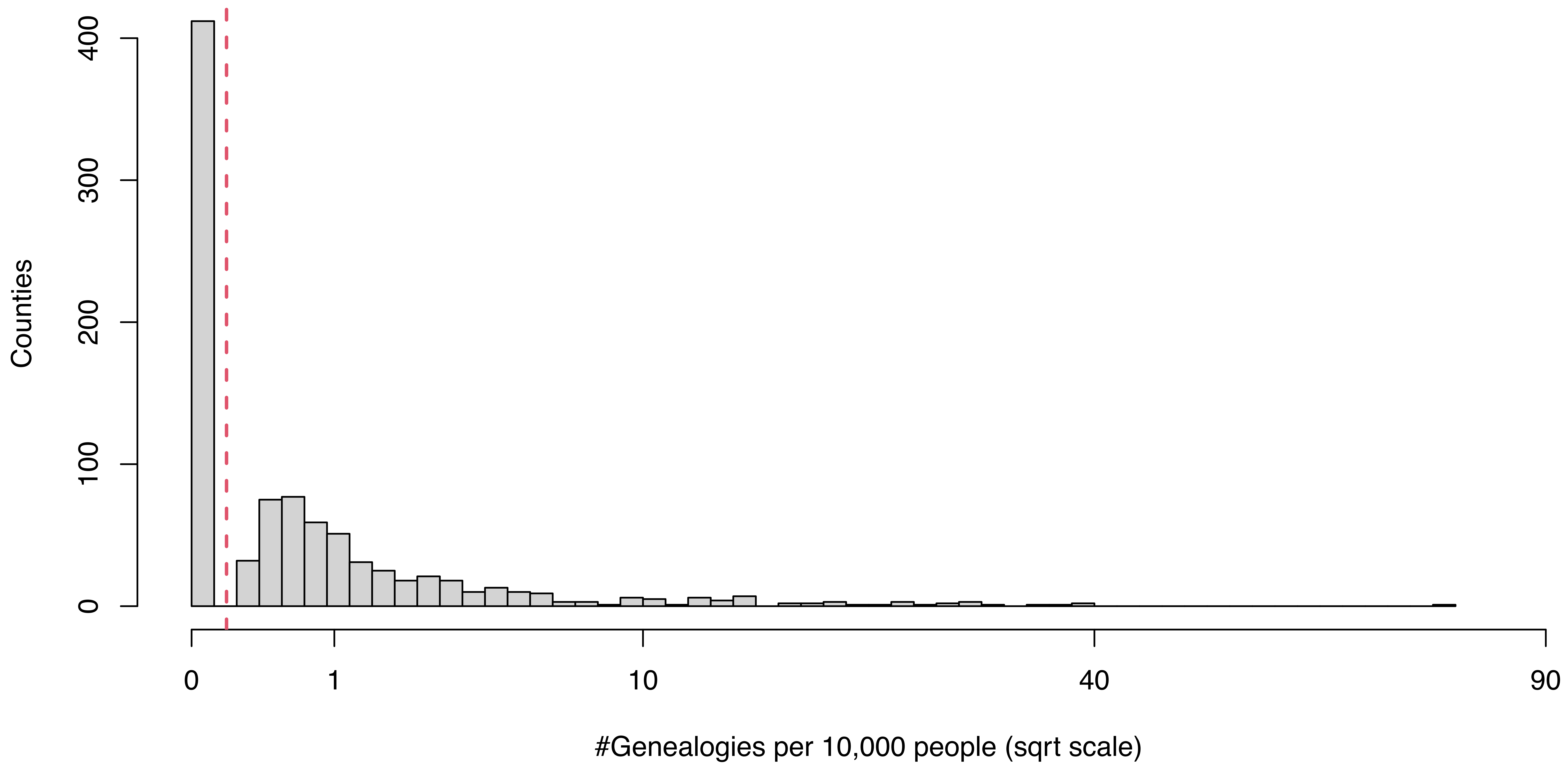
Dynamic Estimates



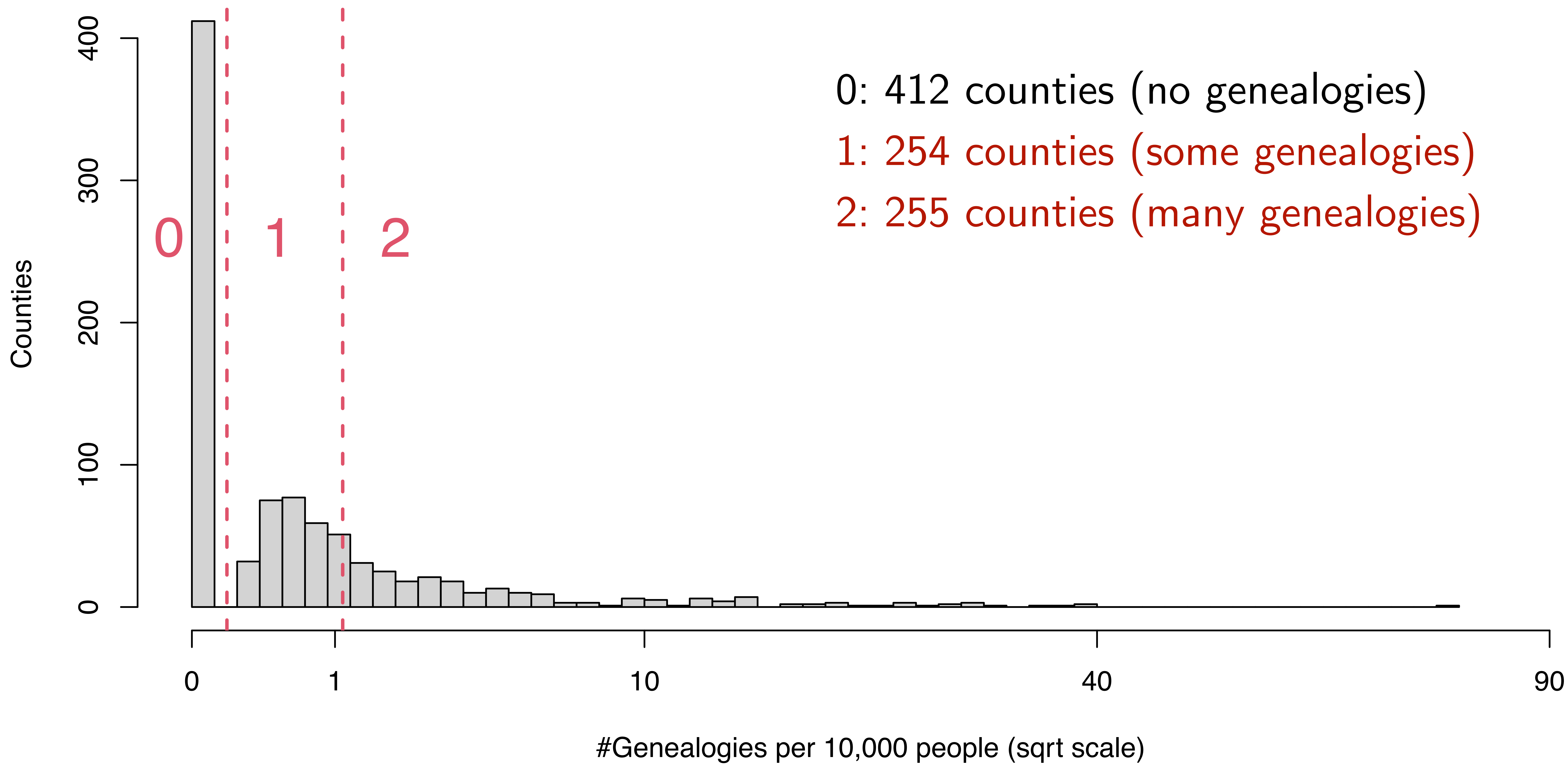
Dynamic Estimates



Multi-valued G



Multi-valued G



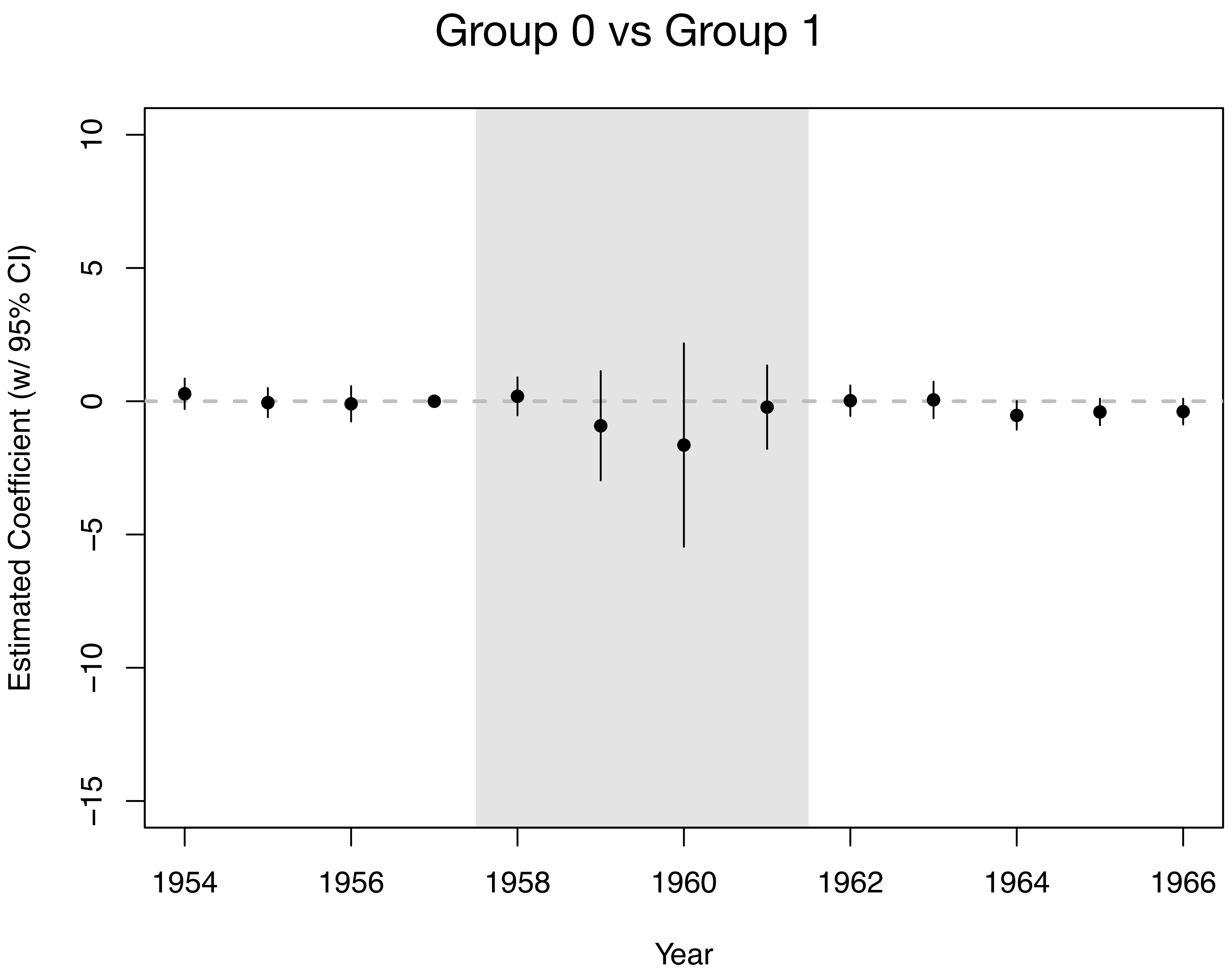
0: 412 counties (no genealogies)

1: 254 counties (some genealogies)

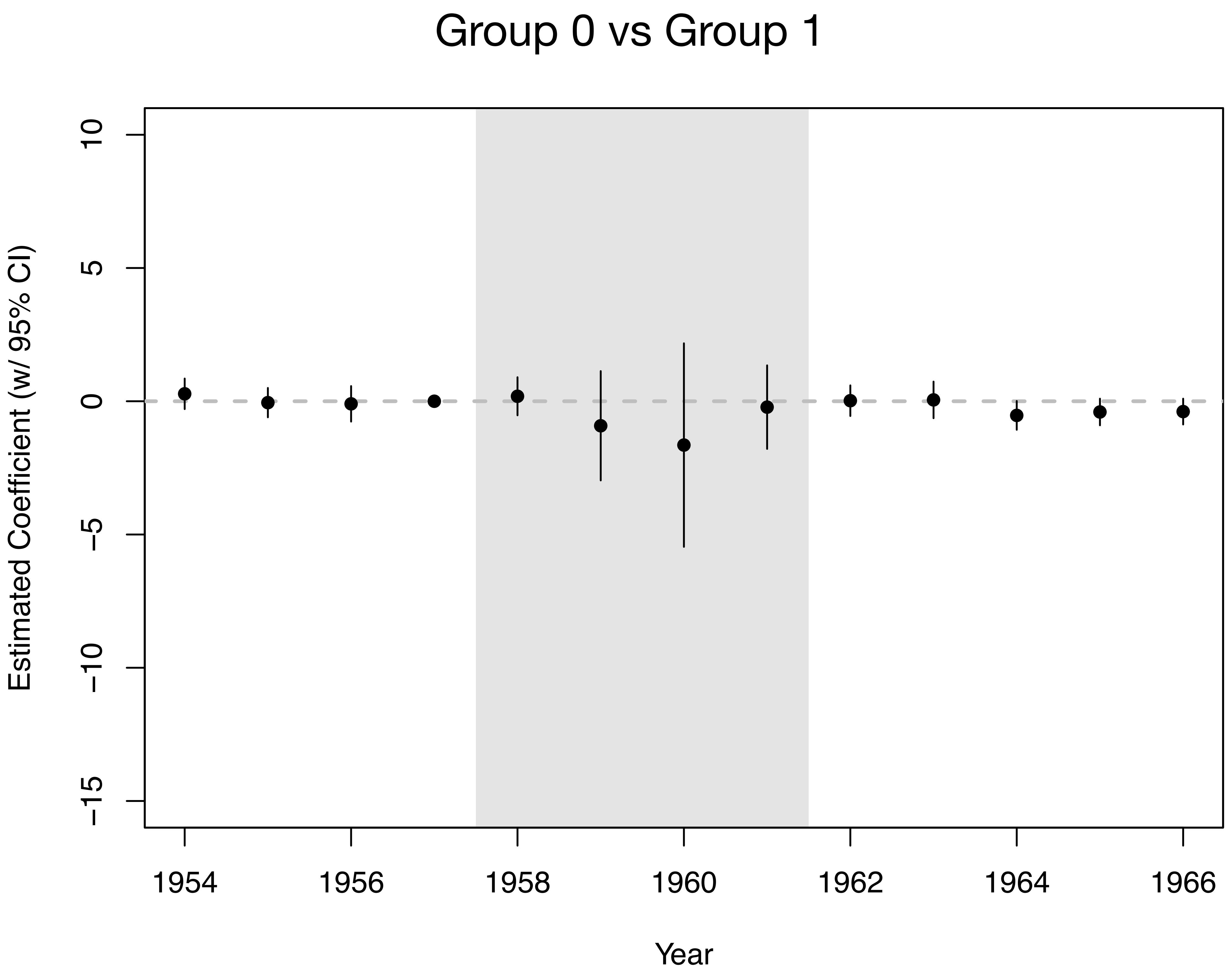
2: 255 counties (many genealogies)

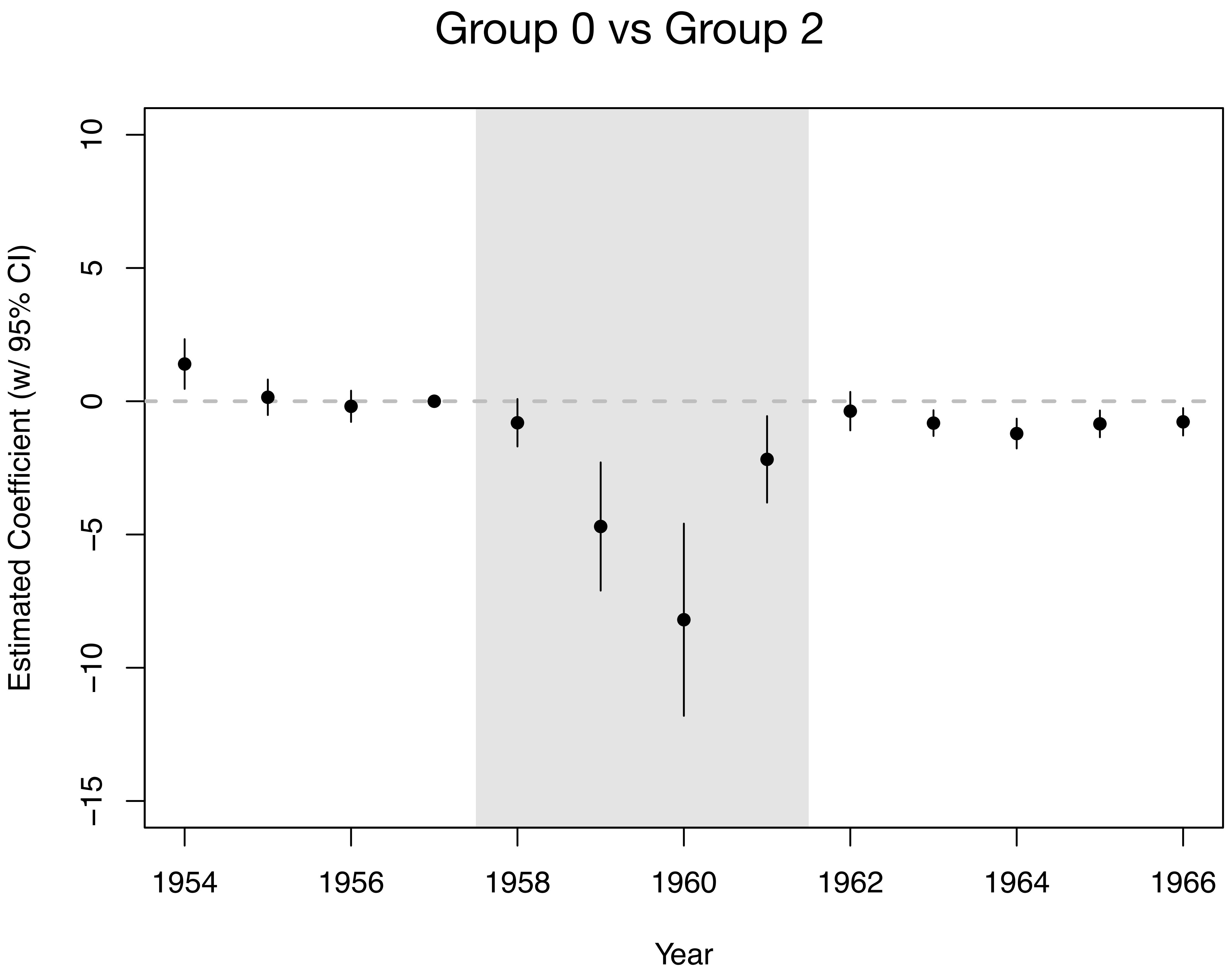
Multi-valued G — AIPW Estimator

Multi-valued G — AIPW Estimator



Multi-valued G — AIPW Estimator





Conclusion & Recommendations

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is <u>no clean control</u> group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies <u>effect modificatio</u>n under NA, PT
  - Identifies <u>causal moderation</u> under NA, PT & ignorability
- What to do in practice

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability
- What to do in practice
  - Assess no anticipation and parallel trends assumptions (by checking pretrends)

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability
- What to do in practice
  - Assess no anticipation and parallel trends assumptions (by checking pretrends)
  - Justify $\Delta Y(g, z) \perp\!\!\!\perp G \mid X$ and show robustness of findings (e.g. by conducting sensitivity analysis)

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability

- What to do in practice
  - Assess no anticipation and parallel trends assumptions (by checking pretrends)
  - Justify $\Delta Y(g, z) \perp\!\!\!\perp G \mid X$ and show robustness of findings (e.g. by conducting sensitivity analysis)
  - If using TWFE models, demean covariates and add interaction terms $G_i \cdot X_i \cdot \text{Post}_t$

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability

- What to do in practice
  - Assess no anticipation and parallel trends assumptions (by checking pretrends)
  - Justify $\Delta Y(g, z) \perp\!\!\!\perp G \mid X$ and show robustness of findings (e.g. by conducting sensitivity analysis)
  - If using TWFE models, demean covariates and add interaction terms $G_i \cdot X_i \cdot \mathrm{Post}_t$
  - Should not automatically assume no carryover effects

Conclusion & Recommendations

- We formulate Factorial DID as a research design with panel data when there is no clean control group
  - Corresponds to a hypothetical factorial experiment w/ group indicator and exposure as two factors
  - Identifies effect modification under NA, PT
  - Identifies causal moderation under NA, PT & ignorability
- What to do in practice
  - Assess no anticipation and parallel trends assumptions (by checking pretrends)
  - Justify $\Delta Y(g,z) \perp\!\!\!\perp G \mid X$ and show robustness of findings (e.g. by conducting sensitivity analysis)
  - If using TWFE models, demean covariates and add interaction terms $G_i \cdot X_i \cdot \mathrm{Post}_t$
  - Should not automatically assume no carryover effects
- Thank you!